



Víctor M. Macías Alemán

## Digitalización y postproducción del patrimonio documental de Fuerteventura

2 al 6 noviembre 2009

## ¿Porqué digitalizar?



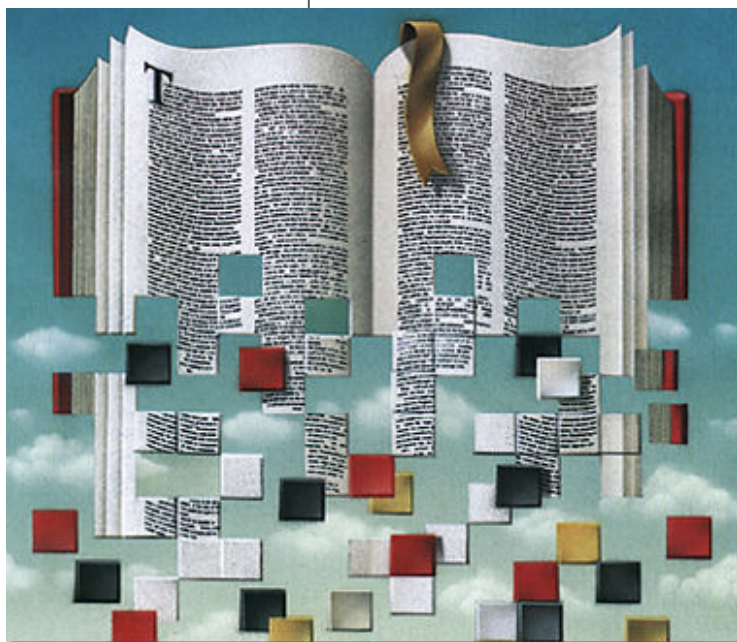
***No mas prototipos, es el momento de la aplicación a gran escala***

Conferencia Internacional sobre digitalización  
del patrimonio cultural europeo (Utrecht, octubre 1999)

- ❖ Preservación de los originales
- ❖ Optimización del espacio
- ❖ Eliminación de los originales innecesarios
- ❖ Difusión de nuestros propios fondos
- ❖ Puesta en valor de contenidos desconocidos
- ❖ Ofrecer nuevos servicios al usuario que a veces este no había siquiera pensado



## ¿Porqué digitalizar?



- ❖ Potenciar el acceso a nuestra cultura por nuevos medios
- ❖ Aumentar la visibilidad de nuestra creación en Internet
- ❖ Estimular el uso de las TIC por la sociedad
- ❖ Favorecer la creación de industrias y empresas de servicios documentales y digitalización
- ❖ Impulsar el desarrollo de nuestras instituciones culturales, educativas e investigadoras.
- ❖ Nuevo medio para favorecer el libre acceso al conocimiento



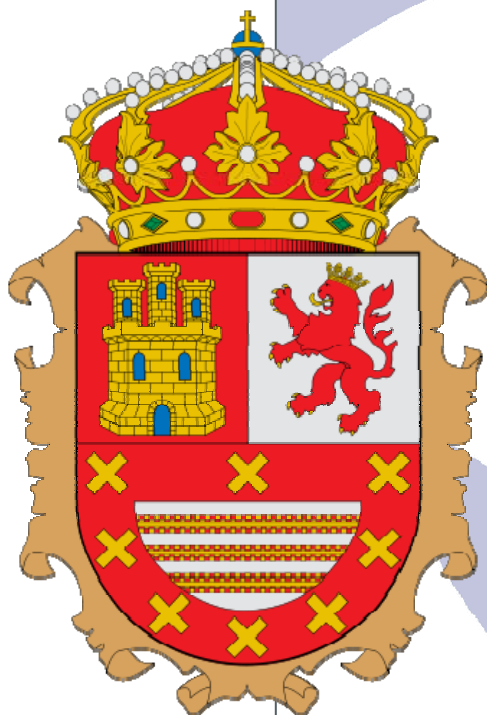
## Más razones para digitalizar



- ❖ Mejorar la gestión documental (conversión analógica a digital)
- ❖ Reducir costes operativos
- ❖ Control de acceso y seguridad (Encriptación, firma electrónica, permisos de acceso, LOPD...)
- ❖ Nuevas oportunidades de negocio
- ❖ Preservación
- ❖ Ampliar el acceso a la información (aumento de puntos de acceso)
- ❖ Difusión y visibilidad de la fondos



## Conocimiento de la tarea. Estudio previo



### ❖ El patrimonio documental de Fuerteventura

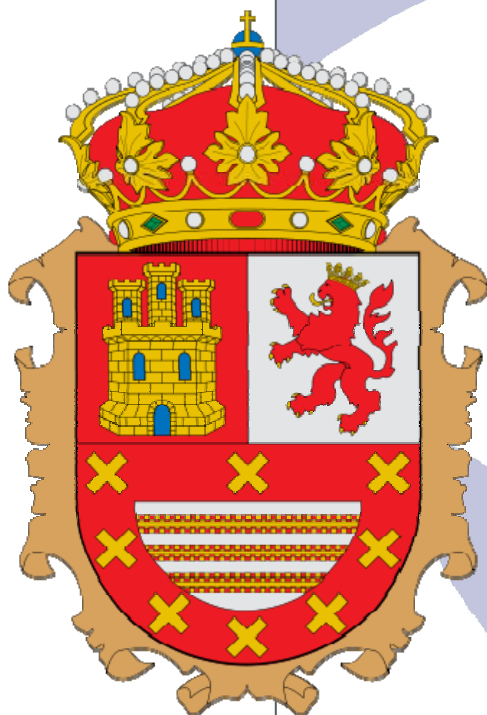
- ❖ Primer impreso específico impreso en Canarias, el *Reglamento de la liga de contribuyentes de Fuerteventura*, de 1883, al que se suma en 1889 el *Resumen de la historia* de Lanzarote y Fuerteventura, de Antonio María Manrique, o los *Estatutos de la Sociedad Anónima Mercantil La Esperanza establecida en Puerto de Cabras* del mismo año, y los estatutos de la Sociedad *El Porvenir*, impreso al año siguiente.

- ❖ En cuanto a las publicaciones periódicas, principia en 1900 con *La Aurora*, editada hasta el año 1905 en Puerto de Cabras aunque impresa en Gran Canaria. Otro intento fue *Erbania*:

*En una imprenta de Puerto del Rosario hallé los plomos de la mancheta de la revista Erbania, que rezaba Semanario de Fuerteventura. Como viñeta alrededor del título se veían unas palmeras y un camello, con el sol nascente al fondo.*

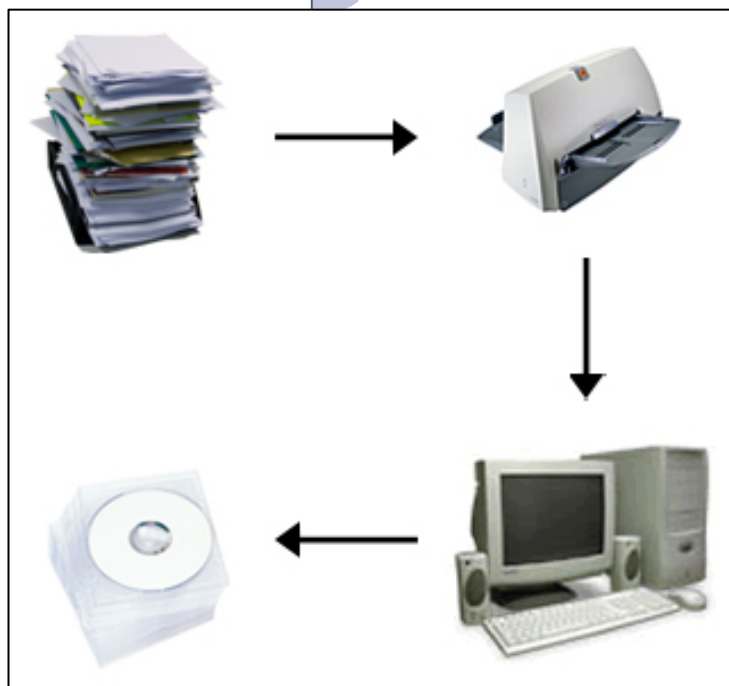
Saavedra Rodríguez, José Antonio. *Catálogo gral. de public. periód. de la prov. de Las Palmas, 1840-1972*. Tesina inédita





## Conocimiento de la tarea. Estudio previo

- ❖ Digitalizada ya *La Aurora* y pendiente de poner en línea para el uso ciudadano.
- ❖ También hay que enlazarse a lo realizado por otras entidades locales o del extranjero [http://humboldt.mpiwg-berlin.mpg.de/hartu\\_diege\\_de\\_01\\_1857-LiSe/](http://humboldt.mpiwg-berlin.mpg.de/hartu_diege_de_01_1857-LiSe/)
- ❖ En general, las tipología documental suele ser la misma en un lado que en otro. Lo que cambian son las circunstancias. En Fuerteventura, como en otros lugares, se debe atender a lo siguiente para rescatar, gestionar, preservar y, en su caso, digitalizar y difundir su patrimonio documental:
  - Archivos y bibliotecas de particulares
  - Archivos y bibliotecas municipales e insulares de la administración pública
  - Archivos y bibliotecas de entidades privadas
  - Archivos y bibliotecas foráneos con documentación sobre Fuerteventura
  - La producción de sus autores y su legado posterior
  - La bibliografía existente sobre Fuerteventura
  - Editoriales y servicios de publicaciones propios
- ❖ También temas de interés extrínseco:
  - Turismo
  - Espacios naturales protegidos
- ❖ ...y no olvidarse del Patrimonio Intangible

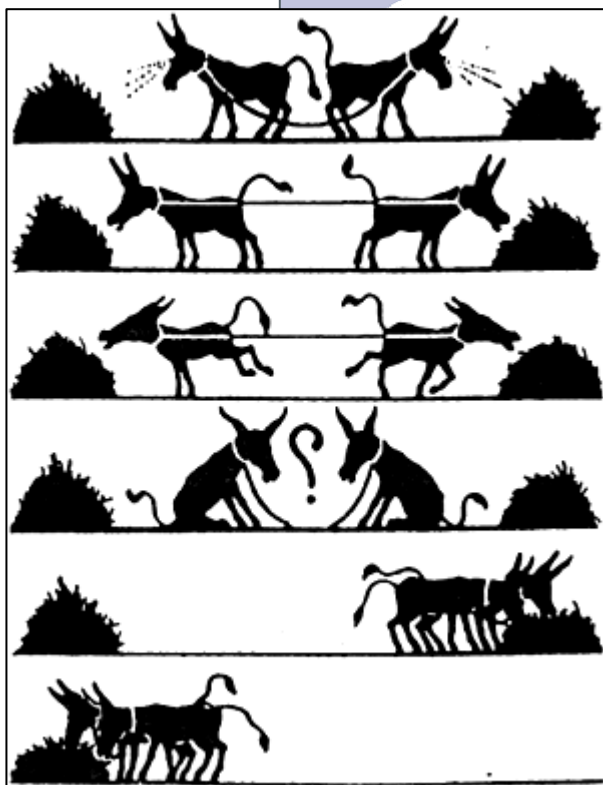


## Fases para acometer la digitalización

- ❖ Análisis
- ❖ Definición
- ❖ Diseño
- ❖ Planificación
- ❖ Ejecución

en cualquier caso con  
una visión práctica  
y teniendo en cuenta que...





**En la digitalización,  
(como en todo)  
siempre es mejor cooperar**

- ❖ Evitamos duplicar esfuerzos
- ❖ Ahorramos costes
- ❖ Digitalizamos más documentación
- ❖ Intercambiamos conocimientos
- ❖ Desarrollamos pautas comunes
- ❖ Acometemos proyectos más ambiciosos
- ❖ Podemos especializarnos en tareas específicas
- ❖ El usuario obtiene más





## Conocimiento de la tarea. Estudio previo

- ❖ Digitalizar nosotros
  - ❖ Conocimientos
  - ❖ Espacio
  - ❖ Hardware+software
  - ❖ Personal
  - ❖ Financiación del trabajo
- ❖ Digitalizar cooperativamente (prensa canaria corriente)
- ❖ Externalizar
  - ❖ Adocat  
<http://www.adocat.com/index.asp?w=pag&s=24&a=72&mn=24>
  - ❖ Cadema Informática  
[http://www.gitdoc.es/ctl\\_arch/3serviesc.htm](http://www.gitdoc.es/ctl_arch/3serviesc.htm)
  - ❖ GCIC. Gestión y custodia de la información Canarias S.L.  
<http://www.gcic.es/gcic/GCIC/published/legal.html>
  - ❖ Tecnodoc. Tecnologías de la documentación, S.A.  
<http://www.tecnodoc.com>
  - ❖ La Tienda de la ULPGC <http://www.tienda.ulpgc.es>



## Estudio de la organización



- ❖ Estructura
- ❖ Funcionamiento interno
- ❖ Agentes que se relacionarán directamente con el desarrollo del proyecto
- ❖ Activos (recursos humanos, infraestructuras, tecnologías y presupuestos)
- ❖ Destacar también los factores y agentes condicionantes
- ❖ Documentos
- ❖ Agentes administradores
- ❖ Usuarios
- ❖ Flujos de información y trabajo



## Estudio de factores externos



- ❖ Datos sobre iniciativas similares (peligro de duplicar esfuerzos)
- ❖ Principales agentes del sector
- ❖ Instituciones que ofrecen pautas, cooperación y financiación
- ❖ Subvenciones
- ❖ Empresas proveedoras de servicios





## Proyecto. Preguntas sobre su misión y objetivos

- ❖ ¿Qué queremos conseguir con la digitalización?
- ❖ ¿Cuáles son los aspectos clave?
- ❖ ¿A quién beneficiará o afectará el proyecto?
- ❖ ¿Qué cambios se pueden producir que afecten a los resultados?
- ❖ Aspectos que habrá que definir en una segunda fase





## Proyecto. Caso supuesto

❖ La Asociación *Islas Afortunadas* es una entidad sin ánimo de lucro que quiere digitalizar 6000 carteles de eventos de carácter regional de todas las épocas. El fondo está controlado mediante una base de datos referencial. Un estudio hecho aconseja la realización de un proyecto de digitalización de este fondo que garantice su preservación y difusión global. Para conseguir financiación pública para llevar a cabo el proyecto, la asociación redacta un proyecto.







## Proyecto. Redacción

- ❖ Texto introductorio de presentación
- ❖ Objetivos estratégicos
- ❖ Estructura
- ❖ Principales funcionalidades
- ❖ Fases de realización
- ❖ Agentes implicados
- ❖ Metodología de trabajo
- ❖ Misión (p. ej., generar un repositorio digital que alojará las imágenes de los carteles consultables por los investigadores y que supondrá también su preservación)





## Proyecto. Redacción

### ❖ Objetivos

- ❖ Realistas (asumir los recursos de que se dispone)
- ❖ Cuantificables
- ❖ Concretos (ceñirse a las acciones programadas)
- ❖ Compatibles (no pueden ser excluyentes entre sí)
- ❖ Sincronizables (calendario)

#### ❖ EJEMPLO:

- ❖ Escoger un sistema que gestione la colección de imágenes digitalizadas y sus metadatos
- ❖ Catalogar el fondo a digitalizar
- ❖ Digitalizar de forma completa el fondo
- ❖ Elaborar una política de preservación
- ❖ ...





## Proyecto. Redacción

### ❖ Coste económico

- ❖ Es difícil de calcular un precio fijo de mercado (diferentes factores a considerar) como
  - ❖ Cantidad documentación
  - ❖ Calidad que se desea
  - ❖ Dimensiones
  - ❖ Soporte de los originales
  - ❖ Estado de presentación
  - ❖ Estado de conservación
  - ❖ Organización del fondo
  - ❖ Postproducción necesaria
  - ❖ Subproductos que se desea obtener
  - ❖ Carga en sistema







## Proyecto. Redacción

### ❖ Coste económico

- ❖ Estimación de precios (margen aprox. 30 %)

Digitalización automática:

- ❖ 1 página A4 b/n 300 ppp TIFF: 0'10 €
- ❖ 1 página A4 grises 8 bits 300 ppp TIFF: 0'20 €
- ❖ 1 página A4 color 24 bits 300 ppp TIFF: 0'30 €

Digitalización manual:

- ❖ 1 página A4 grises 8 bits 300 ppp TIFF: 0'80 €
- ❖ 1 página A4 color 24 bits 300 ppp TIFF: 0'90 €

Postproducción (se añade a c/u 0'09 €) para:

- ❖ Enderezado, visión ajustada, rotación, unión o separación
- ❖ Generación de OCR
- ❖ Generación de PDF multipágina con mejoras





## Proyecto. Redacción

### ❖ Tecnología

- ❖ Analizar los medios técnicos a utilizar
- ❖ No caer en la tentación de ser pioneros
- ❖ No usar tecnologías en desuso u obsoletas

### ❖ Cronología

- ❖ Calendario consensuado y realista
- ❖ Con márgenes para contemplar imprevistos
- ❖ Incorporar penalizaciones por incumplimiento de terceros







## Proyecto. Redacción

❖ El proyecto de organización del fondo de carteles de la Asociación *Islas Afortunadas* tendrá una duración aproximada de 12 meses con la posibilidad de ampliarlo a 13 en el caso de que se produzcan incidencias de carácter técnico que lo justifiquen. Tendrá una dotación presupuestaria de 20.000 € y pretende conseguir ayudas públicas que aportarán un total de 12.000 €. El proyecto incluye las fases de selección documental, captura, tratamiento, alojamiento de las imágenes en un servidor informático de la propia entidad y la elección de un software para gestionar los datos asociados. No se contempla, por el contrario, el diseño y la creación del portal que dará acceso al fondo digitalizado.



## Financiación



### ❖ Propia

#### ❖ Con documentos propios

<http://www.gobiernodecanarias.org/boc/2009/017/013.html>

<http://www.gobiernodecanarias.org/boc/2007/213/019.html>

<http://www.gobiernodecanarias.org/boc/2007/196/009.html>

#### ❖ Con documentos ajenos de nuestro interés

[http://www.boe.es/aeboe/consultas/bases\\_datos/doc.php?coleccion=indilex&id=2008/11171&txtlen=720](http://www.boe.es/aeboe/consultas/bases_datos/doc.php?coleccion=indilex&id=2008/11171&txtlen=720)

## Financiación



- ❖ Externa
  - ❖ Convenios
  - ❖ Subvenciones
    - ❖ Estatales <http://www.boe.es/boe/dias/2009/02/09/pdfs/BOE-A-2009-2234.pdf>
    - ❖ Autonómicas
      - <http://www.gobiernodecanarias.org/boc/2007/033/017.html>
      - <http://sacm.jccm.es/siaci/descripcion.asp?CodigoServicio=IBH>
    - ❖ <http://www.snae.org/ayudas10.es.php>
    - ❖ <http://herencia.net/2008/08/05/subvenciones-para-exposiciones-de-fotografias-antiguas-y-digitalizar-patrimonio-documental/>
  - ❖ Programas de investigación y cooperación
  - ❖ Privada
    - ❖ Patrocinio con publicidad
    - ❖ Mecenazgo
      - ❖ Google Books <http://www.ucm.es/BUUCM/biblioteca/11979.php>



## Identificación, organización y preparación



- ❖ Selección y valoración
  - ❖ Lo más demandado por los usuarios
  - ❖ Lo que se encuentra en peor estado
  - ❖ Lo que deseamos eliminar físicamente
  - ❖ Obras y autores reconocidos
  - ❖ Lo más inaccesible
  - ❖ Lo más raro
  - ❖ Lo más desconocido (pero importante)
  - ❖ Lo más atractivo
  - ❖ Lo que no podemos reeditar de forma impresa
  - ❖ Que tengamos derechos o que hayan prescrito
  - ❖ Cronológico
  - ❖ Topográfico
  - ❖ Temático



## Identificación, organización y preparación



- ❖ Y ahora que está seleccionado...
  - ❖ ¿Estará ya digitalizado?
  - ❖ ¿Existe ya una versión comercial a la venta?
  - ❖ ¿Alguien lo está digitalizando o tiene previsto hacerlo en breve?
  - ❖ ¿Hay versiones intermedias aprovechables (microfilm, impresiones facsímiles...)?





## Identificación, organización y preparación



- ❖ Identificación
  - ❖ Imágenes-fotos
    - ❖ ¿Quién o qué?
    - ❖ ¿Cómo?
    - ❖ ¿Dónde?
    - ❖ ¿Cuándo?
- ❖ Normalización
  - ❖ Imágenes-fotos
    - ❖ ISAD (G) + ISAAR (CPF)
    - ❖ Metadatos EAD



## Identificación, organización y preparación



- ❖ Organización
  - ❖ Respetar el principio de procedencia
  - ❖ Criterios archivísticos
  - ❖ Criterios bibliotecarios
    - ❖ Conocimiento del fondo (expertos temáticos si fuera necesario)
- ❖ Preparación (algunos criterios):
  - ❖ Separar texto (color, grises, b/n) por formatos
  - ❖ Que se puede guillotinar de lo que no
  - ❖ Que se puede trasladar fuera de lo que no
- ❖ Difundir
- ❖ Rentabilizar

## Diseño de la digitalización



### ❖ Fase previa

- ❖ Criterios de selección documental
  - ❖ Derechos de autor y protección de datos personales
  - ❖ Contenido (valor intelectual)
    - ❖ Relevancia que justifique el coste de la digitalización
    - ❖ Representatividad de su periodo
    - ❖ Interés atemporal
    - ❖ Original o copia
    - ❖ Completo o no
- ❖ Demanda
- ❖ Estado físico documentación



## Diseño de la digitalización

### ❖ Fase previa

#### ❖ Preparación de la documentación

- ❖ Inventariado
- ❖ Agrupación
  - ❖ Soporte
  - ❖ Tipología documental
  - ❖ Clasificación orgánica

#### ❖ Tratamiento físico

- ❖ Guillotinar
- ❖ Desgrapar
- ❖ Restauración
- ❖ Duplicados o partes que faltan
- ❖ Ordenación y procesado técnico
- ❖ Expurgado
- ❖ Codificación
- ❖ Ordenación (lotes)
- ❖ Codificación



## Diseño de la digitalización



### ❖ Fase de captura

- ❖ Si es documento con valor legal  
ajustar a la EHA/962/2007  
(BOE 14-4-2007 p. 16451 y ss.)

### ❖ Parametrización

- ❖ Resolución
- ❖ Profundidad de color
- ❖ Rango dinámico
- ❖ Compresión (con o sin pérdida)
- ❖ Formato de salida
- ❖ Peso volumen
- ❖ Crear una tabla con parámetros por colección o tipologías





## Diseño de la digitalización



### ❖ Fase de captura

#### ❖ Indización-Indexación

##### ❖ Automática

- ❖ Códigos de barras

- ❖ OCR / ICR

##### ❖ Manual (posterior a digitalización)

- ❖ CDU

- ❖ Encabezamientos de materia

- ❖ Nomenclatura de la UNESCO



## Diseño de la digitalización



### ❖ Fase de captura

- ❖ Procesamiento gráfico automático
  - ❖ Limpieza aparatos
  - ❖ Calibración
  - ❖ Eliminar sombras, rayas, puntos, dedos, contornos, etc.
  - ❖ Corregir encuadrado (redimensionar), oblicuidad, rotación, unión, separación
  - ❖ Ajuste de luminosidad, contraste, *gamma*
  - ❖ Efecto *moaré* (*descreening*)
  - ❖ Corregir color dominante



## Diseño de la digitalización



### ❖ Fase de captura

- ❖ Inspección de resultados
  - ❖ Doble alimentación
  - ❖ Mala orientación
  - ❖ Incorrecto nombre o numeración
  - ❖ Defectos no corregidos automáticamente





## Diseño de la digitalización

### ❖ Fase de tratamiento

- ❖ Descripción mediante metadatos
  - ❖ Intentemos utilizar estándares, no emplear una descripción *sui-generis*
  - ❖ Metadatos descriptivos (creador, título, fecha, materias...)
  - ❖ Metadatos administrativos (técnicos, de preservación, gestión de derechos...)
  - ❖ Metadatos estructurales (relaciones lógicas y físicas entre las diferentes partes)





## Diseño de la digitalización

### ❖ Fase de tratamiento

- ❖ Esquemas de metadatos (principales):
  - ❖ De propósito general: DCMI (Dublin Core Metadata Initiative) o etiquetas META (HTML)
  - ❖ Para archivos: EAD (Encoded Archival Description)
  - ❖ Para museos: VRA (Virtual Resources Association)
  - ❖ Para la educación: LOM (Learning Object Metadata)
  - ❖ Para ontologías y vocabularios: RDF (Resource Description Framework)
  - ❖ Para gestión de derechos: METS/Rights
  - ❖ Para preservación: PREMIS
  - ❖ Se codifican mediante metalenguajes como XHTML, XML... para su procesamiento informático





## Diseño de la digitalización

### ❖ Fase de tratamiento

#### ❖ Introducción de metadatos

- ❖ Incrustados o internos (EXIF, IPCT...)
- ❖ Externos

\*\*\*\*\*

#### ❖ Creación manual (recomendable descriptivos)

❖ Creación automática (para los de tipo técnico es menos costosa y más eficiente para datos o fondos homogéneos).

Ejs: Nombre del operario o fecha de captura -> S.O. ordenador; resolución, profundidad de color, formato, nombre del fichero, tamaño -> Programa de escaneado, etc.

#### ❖ Importación

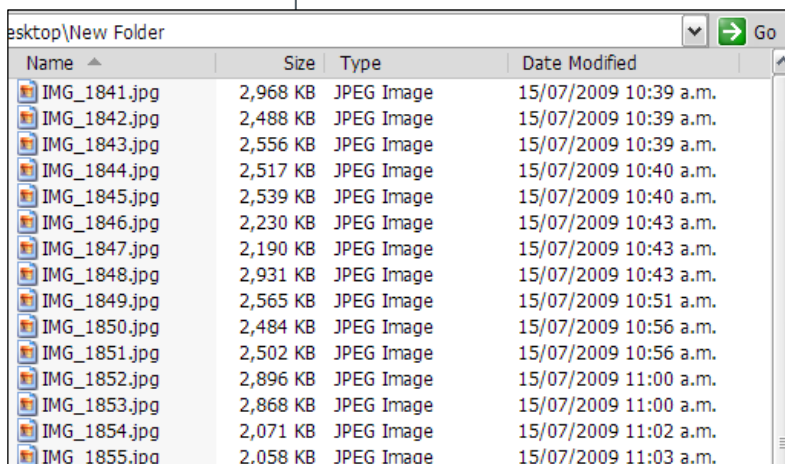


## Diseño de la digitalización

### ❖ Fase de tratamiento

#### ❖ Sintaxis y nomenclatura ficheros

- ❖ Adecuarla a la lógica de la clasificación física del material digitalizado, incluyendo jerarquías.
- ❖ Mejor sólo minúsculas y sin espacios en blanco
- ❖ Evitar cualquier carácter ASCII por encima de ALT 127
- ❖ El nombre del fichero no se debe repetir (error muy común cuando usamos carpetas)
- ❖ Longitud constante con ceros por la izquierda
- ❖ Distintas carpetas para distintos formatos



Name	Size	Type	Date Modified
IMG_1841.jpg	2,968 KB	JPEG Image	15/07/2009 10:39 a.m.
IMG_1842.jpg	2,488 KB	JPEG Image	15/07/2009 10:39 a.m.
IMG_1843.jpg	2,556 KB	JPEG Image	15/07/2009 10:39 a.m.
IMG_1844.jpg	2,517 KB	JPEG Image	15/07/2009 10:40 a.m.
IMG_1845.jpg	2,539 KB	JPEG Image	15/07/2009 10:40 a.m.
IMG_1846.jpg	2,230 KB	JPEG Image	15/07/2009 10:43 a.m.
IMG_1847.jpg	2,190 KB	JPEG Image	15/07/2009 10:43 a.m.
IMG_1848.jpg	2,931 KB	JPEG Image	15/07/2009 10:43 a.m.
IMG_1849.jpg	2,565 KB	JPEG Image	15/07/2009 10:51 a.m.
IMG_1850.jpg	2,484 KB	JPEG Image	15/07/2009 10:56 a.m.
IMG_1851.jpg	2,502 KB	JPEG Image	15/07/2009 10:56 a.m.
IMG_1852.jpg	2,896 KB	JPEG Image	15/07/2009 11:00 a.m.
IMG_1853.jpg	2,868 KB	JPEG Image	15/07/2009 11:00 a.m.
IMG_1854.jpg	2,071 KB	JPEG Image	15/07/2009 11:02 a.m.
IMG_1855.jpg	2,058 KB	JPEG Image	15/07/2009 11:03 a.m.



## Diseño de la digitalización

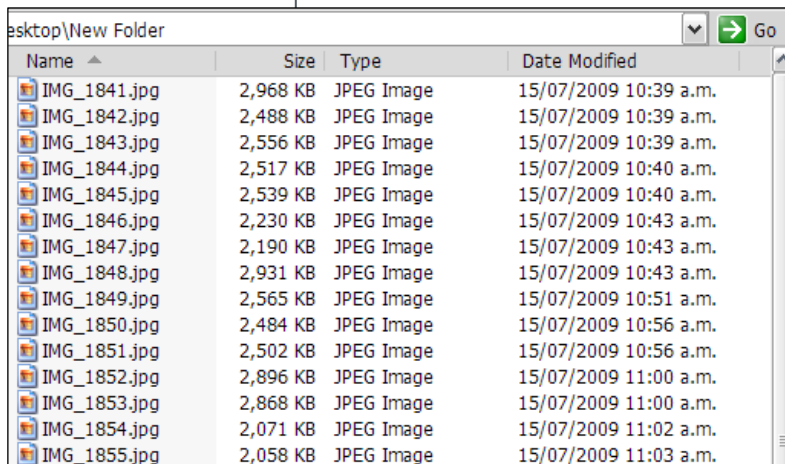
### ❖ Fase de tratamiento

#### ❖ Formato maestro y derivados

- ❖ Versión de alta calidad, no modificado y preservable. Usar estándares, idóneamente TIFF sin comprimir y RAW, WAV, MPEG2, H264

#### ❖ Para derivados hay que saber

- ❖ Uso que se le va a dar
- ❖ Resolución necesaria para transmisión fiel del contenido
- ❖ Dispositivo en el que se va a usar
- ❖ Dónde se va a almacenar
- ❖ Directamente visible navegadores web
- ❖ Ancho de banda de la red a utilizar
- ❖ Necesidad de encriptación, marcas de agua, firmas digitales, etc.
- ❖ Ficheros de mosaico, de previsualización, etc.



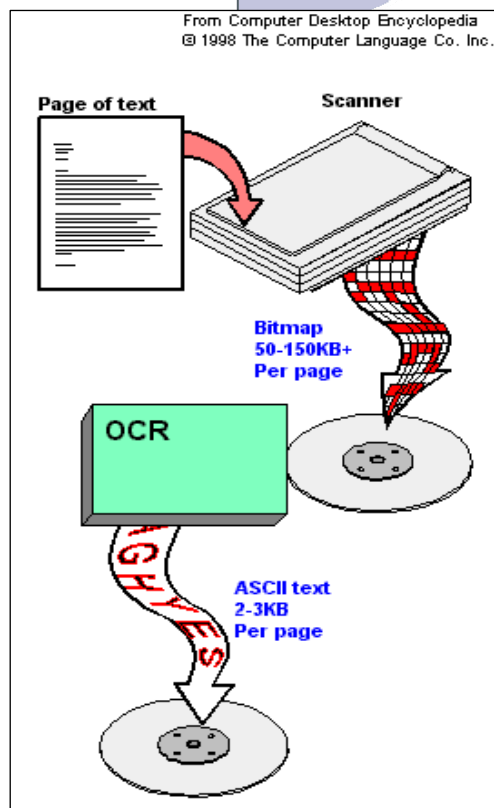
Name	Size	Type	Date Modified
IMG_1841.jpg	2,968 KB	JPEG Image	15/07/2009 10:39 a.m.
IMG_1842.jpg	2,488 KB	JPEG Image	15/07/2009 10:39 a.m.
IMG_1843.jpg	2,556 KB	JPEG Image	15/07/2009 10:39 a.m.
IMG_1844.jpg	2,517 KB	JPEG Image	15/07/2009 10:40 a.m.
IMG_1845.jpg	2,539 KB	JPEG Image	15/07/2009 10:40 a.m.
IMG_1846.jpg	2,230 KB	JPEG Image	15/07/2009 10:43 a.m.
IMG_1847.jpg	2,190 KB	JPEG Image	15/07/2009 10:43 a.m.
IMG_1848.jpg	2,931 KB	JPEG Image	15/07/2009 10:43 a.m.
IMG_1849.jpg	2,565 KB	JPEG Image	15/07/2009 10:51 a.m.
IMG_1850.jpg	2,484 KB	JPEG Image	15/07/2009 10:56 a.m.
IMG_1851.jpg	2,502 KB	JPEG Image	15/07/2009 10:56 a.m.
IMG_1852.jpg	2,896 KB	JPEG Image	15/07/2009 11:00 a.m.
IMG_1853.jpg	2,868 KB	JPEG Image	15/07/2009 11:00 a.m.
IMG_1854.jpg	2,071 KB	JPEG Image	15/07/2009 11:02 a.m.
IMG_1855.jpg	2,058 KB	JPEG Image	15/07/2009 11:03 a.m.



## Diseño de la digitalización

### ❖ Fase de tratamiento

- ❖ Reconocimiento óptico de caracteres
  - ❖ OCR o ICR (mss.) u OMR (marcas)
  - ❖ Para tener el 100% o vectorial o corrección manual
  - ❖ Incorporan correctores ortográficos para cotejar los caracteres con su diccionario
  - ❖ Cuando es necesario emplear vocabularios especializados para el cotejo
  - ❖ Recomendable escanear como mínimo a 300 ppp
  - ❖ Ejecutar el procesado a posteriori por lotes



## Diseño de la digitalización



### ❖ Fase de tratamiento

- ❖ Protección y autenticación
  - ❖ Marcas de agua
  - ❖ Tampones, copyright o sellos digitales
  - ❖ Firma electrónica (ordinaria, avanzada o reconocida. Ley 59/2003 de firma electrónica)
  - ❖ Generación de copias autenticadas
  - ❖ Encriptación de acceso

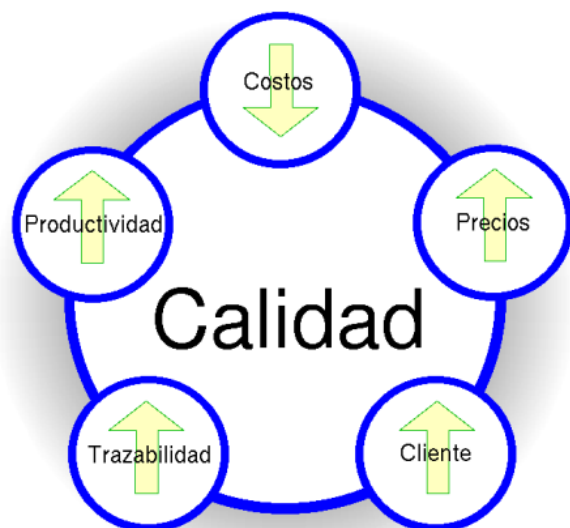


## Diseño de la digitalización

### ❖ Fase de tratamiento

#### ❖ Verificación

- ❖ Establecer muestreo (todo o parte)
- ❖ Maestros y derivados. También entre sí.
- ❖ Fidelidad a los originales
- ❖ Orden y ajustes aplicados
- ❖ Visualización en pantalla
  - ❖ Comprobar a escala 1:1
  - ❖ Patrones/Cartas ajuste color y grises
  - ❖ Histogramas resolución tonal y espacial.
  - ❖ Detección de efectos parásito
- ❖ Impresiones de prueba
- ❖ Estructura y sintaxis de carpetas y ficheros
- ❖ Metadatos generados e introducidos



## Diseño de la digitalización

### ❖ Fase de transferencia y alojamiento

#### ❖ Preservación

- ❖ Definir una política detallada
- ❖ Metadatos PREMIS
- ❖ Soportes fiables y no fiables
- ❖ Redundancia

#### ❖ Protocolos y lenguajes estándar

- ❖ Z39.50
- ❖ OAI-PMH
  - ❖ Proveedor de datos
  - ❖ Proveedor de servicios (cosechador)



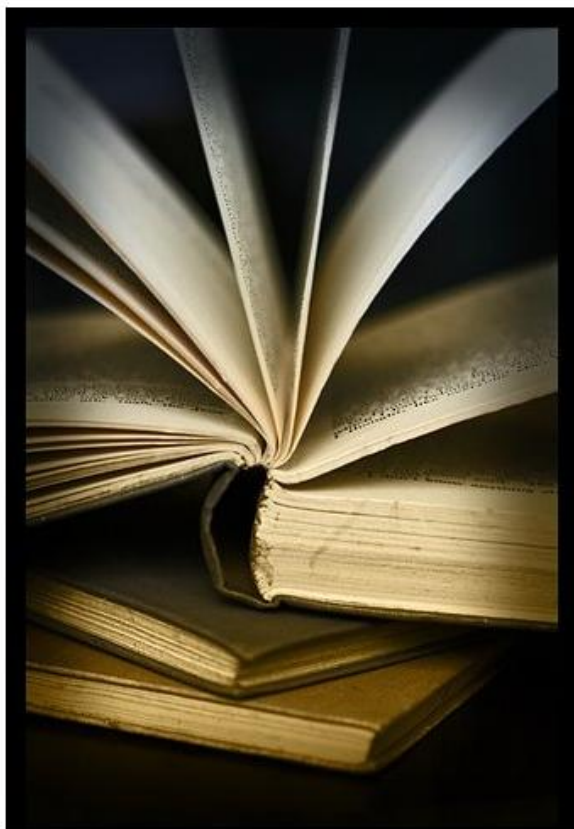


## Diseño de la digitalización



### ❖ Fase de transferencia y alojamiento

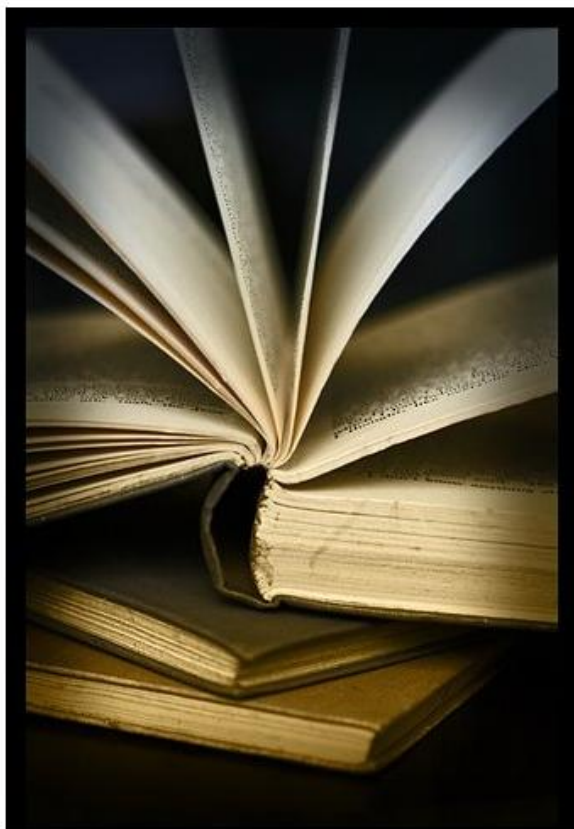
- ❖ Recuerda que van a mirar si tienes o usas...
  - ❖ Z39.50 (opcional)
  - ❖ OAI-PMH (obligatorio)
  - ❖ DTD-CSS-XSL
  - ❖ PURL / URI
  - ❖ OpenURL
  - ❖ XML



## Tipología documental

- ❖ Soportes analógicos:
  - ❖ Presentación: texto, imagen, audio o vídeo.
  - ❖ Material: papel, cartón, plástico, vidrio, cinta magnética, disco microsurco, celuloide, etc.
  - ❖ Conservación: flexible, rígido, deteriorado
  - ❖ Formato: encuadernado, suelto, plegado, multiparte, enrollado...
  - ❖ Dimensiones: desde pocos centímetros hasta A0 o más...
- ❖ Procedencia: propia o ajena
- ❖ Origen: Inédito (manuscrito) o publicado
- ❖ Lugar de realización del trabajo
- ❖ Condicionantes: organizada, no organizada, completa, incompleta, pública, confidencial, varios ejemplares o ejemplar único...





## Tipología documental

- ❖ Soportes digitales:
  - ❖ Material: disco óptico, cinta magnética, disquetes, en línea
  - ❖ Sistema operativo: Linux, Windows, Mac, otros...
  - ❖ Software necesario Latex, Wordperfect, Word, Keynote, Powerpoint, PDF, DjVU...



## Digitalización



- ❖ Equipamiento técnico
  - ❖ Hardware
    - ❖ Servidores
    - ❖ Estaciones de trabajo
    - ❖ Escáneres
    - ❖ Otros dispositivos de captura (toca discos, etc.)



## Digitalización



- ❖ Equipamiento humano (siempre personal especializado)
  - ❖ Bibliotecarios o archiveros
  - ❖ Informáticos
  - ❖ Proveedores documentación
  - ❖ Informantes
- .....
- ❖ BULPGC:
  - ❖ 4 bibliotecarios
  - ❖ Apoyo parcial SIC
  - ❖ Apoyo empresas mantenimiento y desarrollo
  - ❖ Taller de empleo y Escuelas Taller.





## Digitalización



- ❖ Equipamiento humano.  
Nuestro perfil para:
  - ❖ Analizar nuestras necesidades
  - ❖ Analizar y conseguir recursos (humanos y técnicos)
  - ❖ Conocer el sector (soft, hard, normas y pautas, protocolos, usuarios, sector empresarial para mantenimiento...)
  - ❖ Actuar como gestores globales
  - ❖ Afrontar problemas informáticos. Experimentar (voz de texto, PDX...)
  - ❖ Relacionarnos (contactar, viajar, preguntar, conseguir nuevos fondos)



## Digitalización



- ❖ Tipos de escaner y dispositivos de captura para:
  - ❖ Textos-papel
  - ❖ Planos y cartelería
  - ❖ Microfilm
  - ❖ Fotografías e imágenes en papel
  - ❖ Fotografías e imágenes en placas, acetato, etc.
  - ❖ Cinta (casete, DAT, MiniDV, VHS)
  - ❖ Discos vinilo o magnéticos (Minidisc)
  - ❖ ...





## Digitalización

- ❖ Lo que tenemos que tener en cuenta a la hora de elegir (escáner o cámara):
  - ❖ Área de exploración (A4, A3...A0)
  - ❖ Profundidad (bits)
  - ❖ Resolución tonal/color
  - ❖ Resolución óptica (ppp) (nunca “interpolados”)
  - ❖ Zoom óptico (cámaras), nunca “zoom digital”
  - ❖ Uniformidad de la captura
  - ❖ Velocidad de captura (págs/min. a X resol. y X bits)
  - ❖ Automatización de procesos y adaptación al soporte documental
  - ❖ Transferencia de datos USB 2.0
  - ❖ Generación de metadatos
  - ❖ Coste
  - ❖ Mantenimiento y asistencia técnica



## Digitalización

### ❖ Cuando elegimos una cámara fotográfica digital:

- ❖ Calidad (Nikon, Canon, Pentax, Olympus, Panasonic-Leica...)
- ❖ Tipo reflex (objetivos intercambiables)
- ❖ Autoenfoco con posibilidad de modo manual
- ❖ Resolución (actualmente no menos de 12 Megapixels)
- ❖ Memoria tarjetas tipo flash SD/MiniSD/MicroSD
- ❖ Formatos de salida de la imagen RAW y JPEG
- ❖ Exposición y obturación con múltiples medidas y uso manual
- ❖ Transferencia USB 2.0
- ❖ Zoom óptico de al menos 3x (nunca “zoom digital”)
- ❖ Pantalla LCD de al menos 2'5" y buen contraste
- ❖ Interface TWAIN manejable desde ordenador
- ❖ Disparador externo opcional
- ❖ Montable en trípode
- ❖ ... y además una buena mesa de reproducción con iluminación



## Digitalización



### ❖ Escáneres. Marcas

- ❖ Epson <http://www.epson.es/cgi-bin/epson-es/showlistingpage.cgi?typo=web&search=scandocum>
- ❖ Fujitsu <http://www.fujitsu.com/es/services/products/peripherals/scanners/>
- ❖ Canon [http://www.canon.es/For\\_Work/Products/Document\\_Imaging\\_Systems/High\\_Speed\\_Document\\_Scanners/index.asp](http://www.canon.es/For_Work/Products/Document_Imaging_Systems/High_Speed_Document_Scanners/index.asp)
- ❖ SMA <http://www.sma-edoc.com>
- ❖ Kodak [http://graphics.kodak.com/docimaging/ES/es/Products/Document\\_Scanners/index.htm](http://graphics.kodak.com/docimaging/ES/es/Products/Document_Scanners/index.htm)
- ❖ Atiz <http://www.atiz.com>
- ❖ Zeutschel <http://www.zeutschel.com>





## Digitalización

Tabla: Comparación de escáneres

Tipo de escáner	Plano	Con alimentador de hojas	De tambor	Para diapositivas/película	Para microfilm	Cámara digital
Tecnología del sensor*	CCD/CIS	CCD	PMT	CCD	CCD	CCD/CMOS
Tipos de medios**	R, T	R	R, T	T	T	R, T
Rango tonal***	B/W, G, C	B/W, G	G, C	G, C	B/W, G, C	G, C
Limitaciones de tamaño típicas (pulgadas)	11x17 8,5x14 w/ADH 8,5x11 gama baja	8,5x14 de mesa 24x36 de pie	12x17	35mm, 4x5	35mm. Algunos manejan 16 mm y 105 mm, fichas	Ninguna
Margen de resolución (dpi)	300-1200	200-400	1200-8000	1000-5000	5000-8500	Arreglo de píxeles: 1600x1200— 12000x10000
Margen de precio (US\$)	100-50.000+	5.000-100.000+	10.000-100.000+	500-30.000	50.000-250.000.	200-40.000
Uso típico	hojas simples	Documentos de negocios uniformes, documentos, tamaño extra grande	pre-impresión	Diapositivas originales, intermedios	Película, fichas, tarjetas de apertura	Volúmenes encuadernados, material en 3-D, material frágil

\* CCD = Dispositivos Acoplados de Carga; CIS = Sensor de Imagen de Contacto; PMT = Tubo Fotomultiplicador; CMOS = Semiconductor de Óxido de Metal Complementario

\*\* R = Reflectivo; T = Transparente

\*\*\* B/W = Blanco y Negro; G = Escala de Grises; C = Color

\*\*\*\*ADH = Sujetador Automático de Documentos.

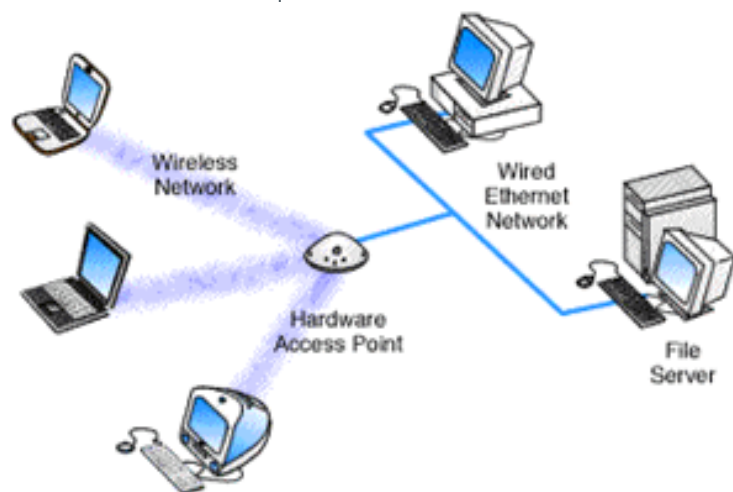
## Digitalización



- ❖ Equipamiento técnico
  - ❖ Hardware
    - ❖ Los escáneres de la BULPGC
      - ❖ Canon MS-300
      - ❖ Minolta PS7000
      - ❖ Fujitsu M4097D
      - ❖ Fujitsu fi5750C
      - ❖ Nikon 5000ED
      - ❖ Nikon 4000ED
      - ❖ Epson GT15000
      - ❖ Epson GT10000
      - ❖ Epson 1640xl
      - ❖ Epson GT2500



## Digitalización



### ❖ Equipamiento técnico

#### ❖ Hardware

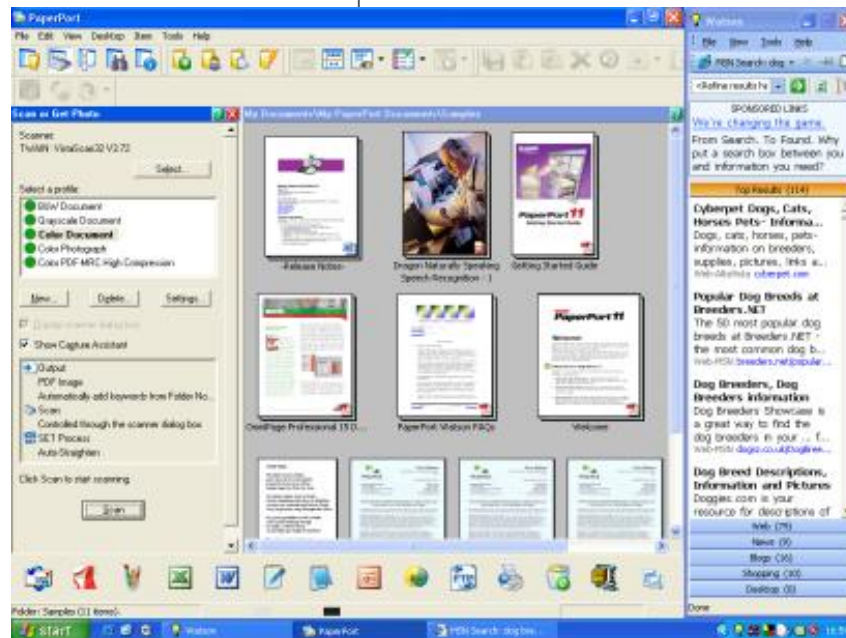
- ❖ PCs de alta capacidad interconectados
- ❖ Escáneres
- ❖ Otros dispositivos de captura
- ❖ Red de comunicaciones Gigabyte
- ❖ Hosting
- ❖ Multiredundancia (“alta disponibilidad”)

## Digitalización

### ❖ Software

#### ❖ Digitalización y procesado (BULPGC)

- ❖ Interface escaner TWAIN (Captura)
- ❖ Paperport (Captura-postproceso)
- ❖ Omnipage (Postproceso)
- ❖ Adobe Acrobat (Captura-postproceso)
- ❖ Irfanview (Postproceso)
- ❖ Windows/Total Commander (Postproceso)
- ❖ Super (Postproceso)
- ❖ I'M too DVD ripper (Postproceso)
- ❖ Adobe Audition (Postproceso)
- ❖ DVD Shrink (Postproceso)
- .....
- ❖ Photoshop, The Gimp, etc. (Postproceso)



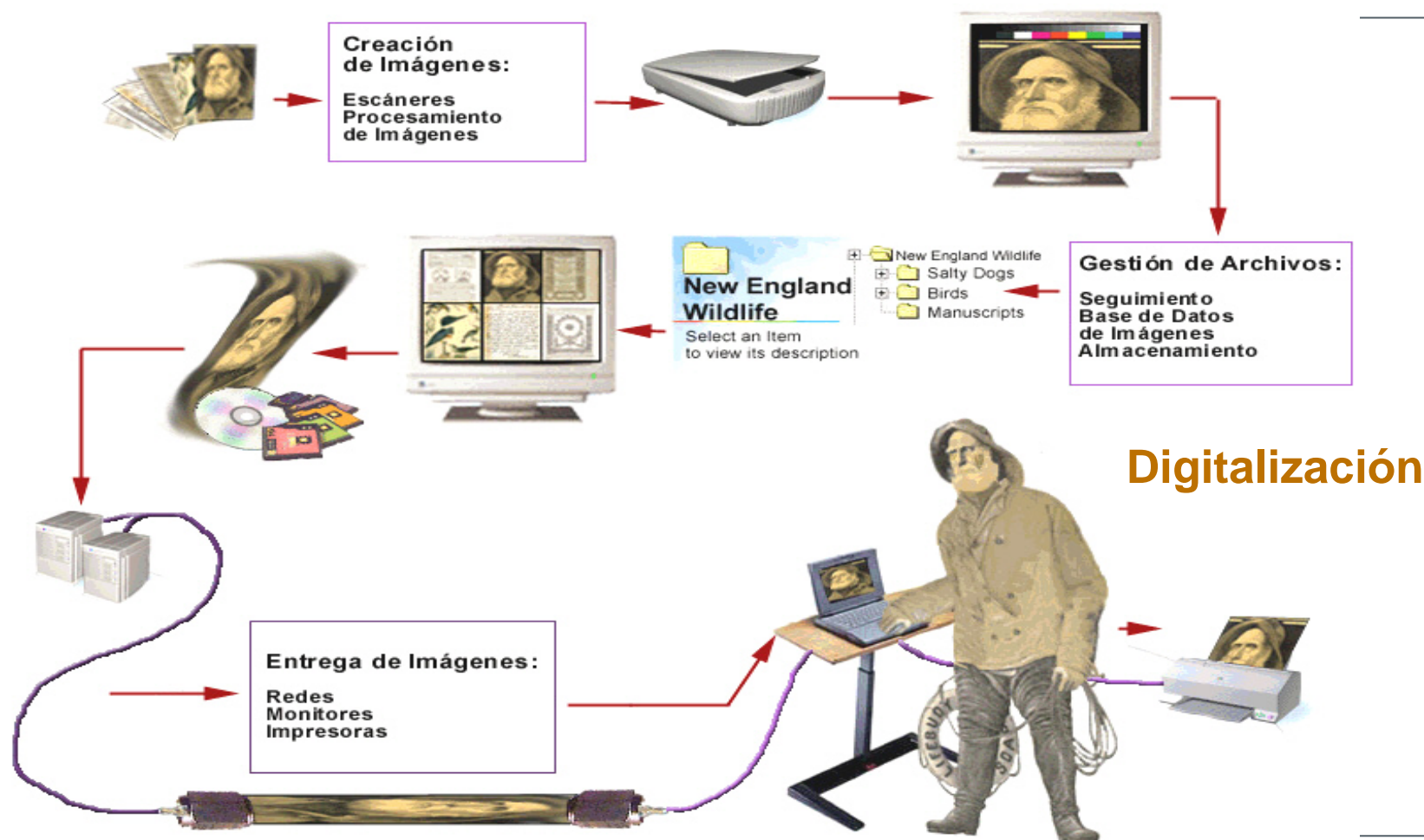
## Digitalización



- ❖ Almacenamiento
  - ❖ Discos ópticos
  - ❖ Duplicadoras de disco
  - ❖ Discos duros portátiles
  - ❖ Discos duros en ordenador personal
  - ❖ Discos duros en servidor (RAID+cinta)
  - ❖ SAN/NAS
- ❖ Valorar
  - ❖ Velocidad de lectura, transferencia y escritura
  - ❖ Capacidad
  - ❖ Fiabilidad
  - ❖ Estandarización
  - ❖ Coste



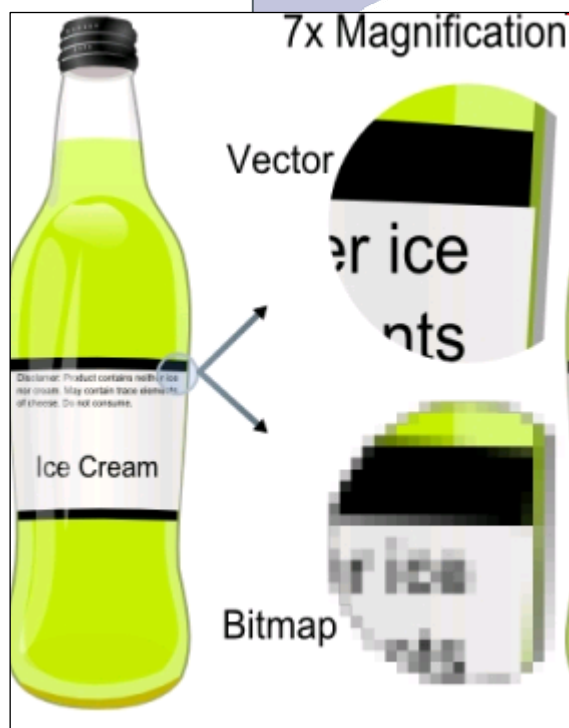








## Digitalización



### ❖ Imágenes Raster-Bitmap vs Vectoriales

- ❖ Las imágenes de barrido o raster son más adecuadas para representar documentos de texto o imágenes reales.
- ❖ Las imágenes vectoriales son más adecuadas para representar planos técnicos y figuras geométricas.
- ❖ Las imágenes de barrido no son escalables sin pérdida de calidad, es decir, pierden calidad a medida que se va haciendo zoom sobre ellas, algo que no sucede con las imágenes vectoriales.
- ❖ Normalmente (\*), una imagen raster se crea con un digitalizador a partir de un original, mientras que una imagen vectorial se crea directamente desde una aplicación específica para ello.
- ❖ Resumiendo: siempre más calidad y menos peso

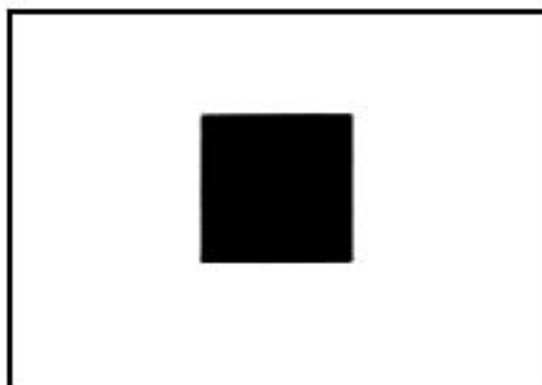
## Digitalización

### ❖ Parámetros para la digitalización

#### ❖ Profundidad

- ❖ 1 bit (1 tono, blanco y negro)
- ❖ 8 bits (256 tonos, escala de grises)
- ❖ 24 bits (16 millones de tonos, color verdadero)

1	1	1	1	1	1	1	1	1	1
1	0	0	0	1	1	0	0	0	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	0	0	0	0	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	0	0	0	1	1	0	0	0	1
1	1	1	1	1	1	1	1	1	1



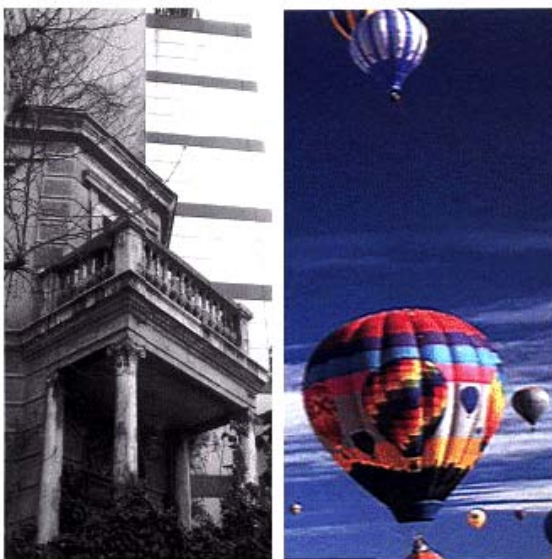
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	0	0	0	1	1	1
1	1	1	1	0	0	0	1	1	1
1	1	1	1	0	0	0	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1





## Digitalización

*ickCam Traveler  
su versatilidad,  
característica que la  
vertido en la prim  
ara de Logitech c  
cionar de modo di  
o cámara digital  
noma y como We  
l de utilizar y con  
rete de software  
es posible casi to  
ara nos ha demo  
las nuevas tecnol  
iversión no tienen  
los factores  
rapuestos.*



- ❖ Parámetros para la digitalización
  - ❖ Profundidad
    - ❖ 1 bit (1 tono, blanco y negro)
    - ❖ 8 bits (256 tonos, escala de grises)
    - ❖ 24 bits (16 millones de tonos, color verdadero)
  - ❖ Fidelidad al original



## Digitalización

- ❖ Parámetros para la digitalización

- ❖ Resolución

- ❖ Se expresa en puntos por pulgada (ppp / dpi)
    - ❖ 300 ppp debe ser el mínimo

- ❖ Pautas adoptadas variables

- ❖ USA

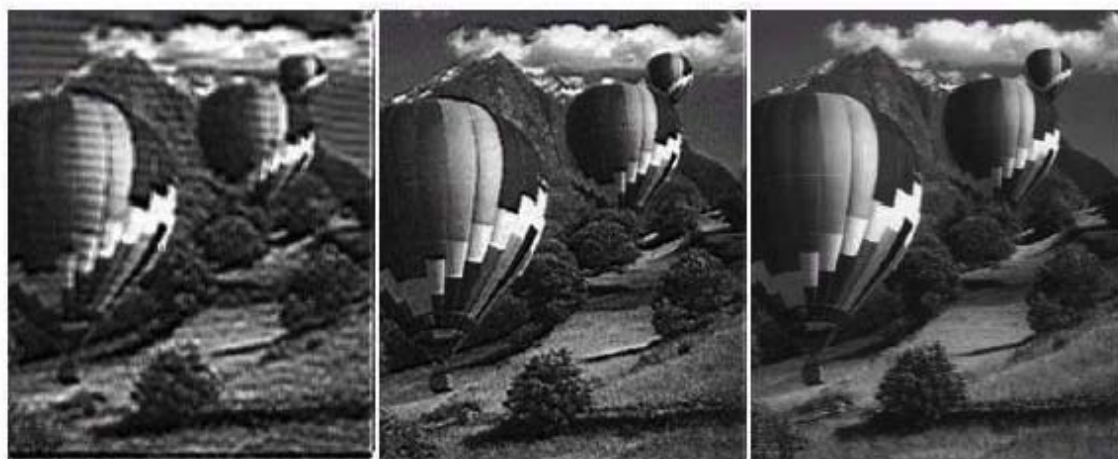
- <http://www.library.cornell.edu/preservation/tutorial-spanish/presentation/table7-2.html>

- ❖ CBUC

- [http://docs.cbuc.cat/fitxers/5digital/Public0507Minims\\_Spa.pdf](http://docs.cbuc.cat/fitxers/5digital/Public0507Minims_Spa.pdf)

- ❖ FEDAC

- <http://www.fotosantiguascanarias.org/digitalizacionpt.php>



Ejemplos a 100, 200 y 300 ppp



## Digitalización

Documento impreso textual sin gráficos o con gráficos simples	Original B/N: 300 ppp, 1 bit (2 tonos)
Manuscritos	Original B/N: 300 ppp, 8 bits (256 tonos) Original color: 300 ppp, 24 bits (color verdadero)
Documento impreso textual con gráficos detallados o fotografías	
Fotografías	
Mapas	
Carteles	
Diapositivas y negativos de 35 mm.	Original B/N: 2048 x 3072 pixels, 8 bits (256 ton.) Original color: 2048 x 3072 pixels, 24 bits (color verdad.)

Para documentos inferiores a A4, como A5, se utilizará como mínimo la siguiente resolución:

Documento impreso textual sin gráficos o con gráficos simples	Original B/N: 600 ppp, 1 bit (2 tonos)
Manuscritos	Original B/N: 600 ppp, 8 bits (256 tonos) Original color: 600 ppp, 24 bits (color verdadero)
Documento impreso textual con gráficos detallados o fotografías	
Fotografías	
Mapas	

Para documentos inferiores al tamaño A5 se estudiará caso por caso la resolución más adecuada.

Pautas adoptadas en la BULPGC para la digitalización de texto e imágenes

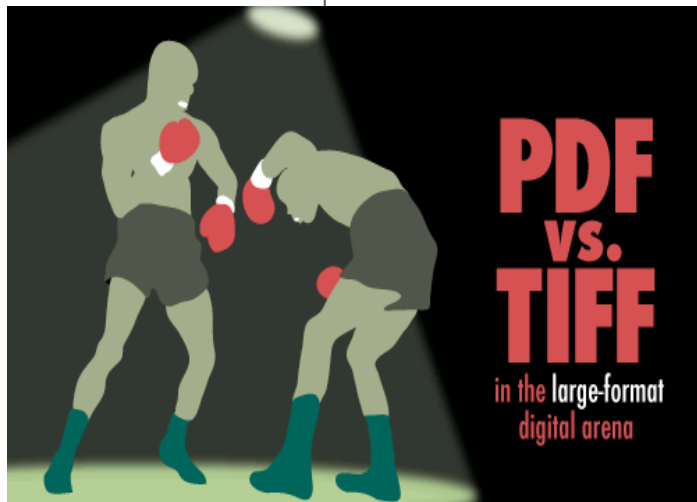


## Digitalización

### ❖ Ficheros y calidades en la BULPGC (mínimos):

#### ❖ Texto e imágenes

- ❖ TIFF sin compresión, 300 ppp (diapositivas 2400 ppp) siempre facsimilar (b/n, grises o color según corresponda), unión o separación de imágenes, etc.
- ❖ TXT con codificación
- ❖ JPEG desde TIFF intermedio redimensionado, compresión calidad 80%, 1'5 Mb max. planos, 700 Kb max. Resto
- ❖ PDF (sólo para documentos textuales) desde TIFF intermedio redimensionado, OCR incrustado, miniaturas, marcadores, visualización a página completa, sin seguridad, con metadatos insertados, compresión JBIG-2 y 75% para JPEG color, visualización a una página, copyright.



## Digitalización

### ❖ Ficheros y calidades en la BULPGC (mínimos):

#### ❖ Audio

- ❖ WAV. Códec PCM sin compresión, 44'1 Khz de muestreo, 16 bit estéreo.
- ❖ FLV (Flash)

#### ❖ Vídeo

- ❖ MPEG-2 a 720x575 pixels, PAL, formato igual al original 4:3 ó 16:9
- ❖ FLV (Flash)

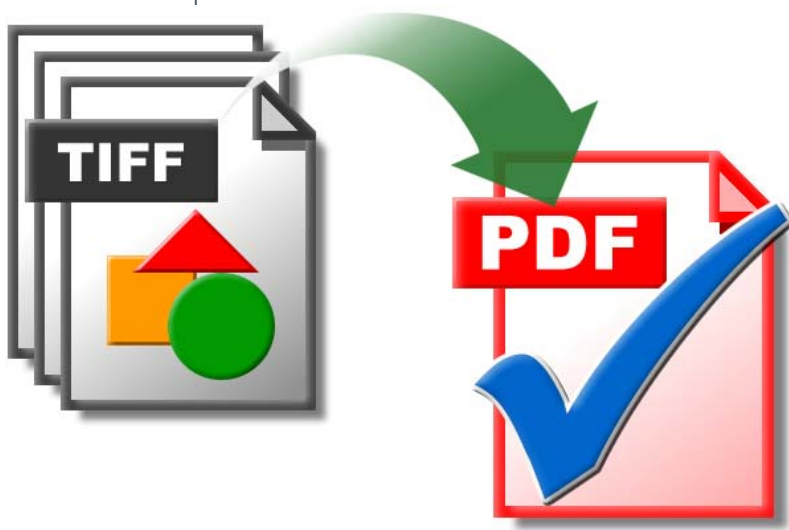
#### ❖ Soportes físicos de entrega

- ❖ CD-R 700 Mb sin multisesión
- ❖ DVD-R 4'7 Gb sin multisesión
- ❖ Discos duros externos USB 2.0





## Digitalización



### ❖ Parámetros para la digitalización

#### ❖ Formatos (texto e imagen)

<http://www.library.cornell.edu/preservation/tutorial-spanish/presentation/table7-1.html>

- ❖ TIFF
- ❖ JPEG
- ❖ PDF
- ❖ DJVU <http://djvu.org>

#### ❖ Formatos (audio y vídeo)

- ❖ MPEG-2
- ❖ MPEG-4
- ❖ WAV
- ❖ MP3
- ❖ Flash
- ❖ Real
- ❖ Quicktime

## Digitalización

### ❖ Otros formatos emergentes...

❖ **PDF / A / X / E**. El más utilizado debido a su fácil uso en el enorme parque instalado de PCs, ordenadores portátiles y netbooks.

❖ **Mobipocket** (.prc, .mobi)

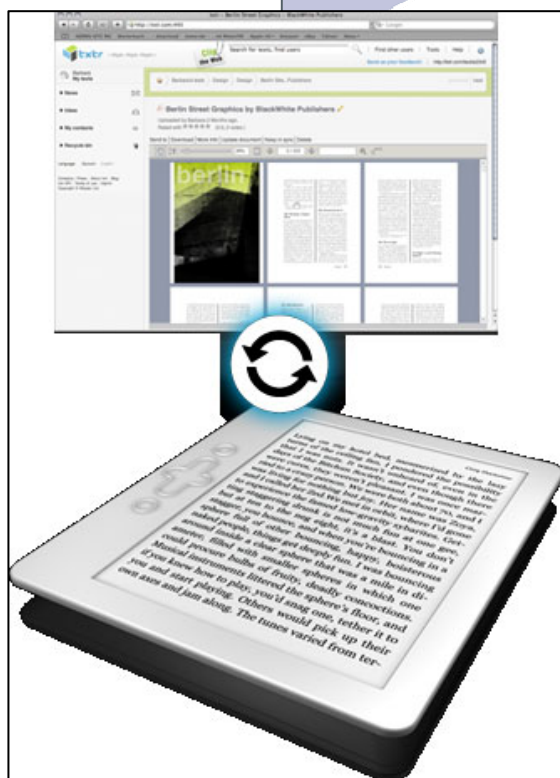
❖ **Amazon Kindle** (.azw). Está basado en Mobipocket.

❖ **eReader** (.pdb), de Palm Digital Media, con un buen soporte por parte de smartphones y pdas,

❖ **Microsoft Reader** (.lit),

❖ **EPUB**. Orientado al uso en dispositivos de tinta electrónica, es de reciente creación (2007) y junto con PDF es el único con soporte internacional

❖ **FictionBook**



Tipo de Imagen	Profundidad Bits/pixel	Resolución Puntos/pulgada	Tamaño Archivo MegaBytes
Blanco & Negro	1	300	1,09
Blanco & Negro	1	400	1,93
16 Grises	4	300	4,36
16 Grises	4	400	7,72
256 Grises	8	300	8,72
256 Grises	8	400	15,44
1.024 Grises	10	300	10,90
1.024 Grises	10	400	19,30
4.096 Grises	12	300	13,08
4.096 Grises	12	400	23,16
8 Colores	3	300	3,27
8 Colores	3	400	3,72
256 Colores	8	300	8,72
256 Colores	8	400	15,44
65.536 Colores	16	300	17,44
65.536 Colores	16	400	30,88
16.777.216 Colores	24	300	26,16
16.777.216 Colores	24	400	46,32
4.294.967.296 Colores	32	300	34,88
4.294.967.296 Colores	32	400	61,76

## Digitalización

❖ Tamaño de los resultados

Cuadro comparativo  
sin compresión para A4



## Diseño de la digitalización

### ❖ Peso - volumen



#### *Fórmula 1 para el tamaño de archivo*

**Tamaño de archivo = (altura x ancho x  
profundidad de bits x dpi<sup>2</sup>) / 8**

Si se proporcionan las dimensiones de píxel, multiplíquelas entre sí y por la profundidad de bit para determinar la cantidad de bits presentes en un archivo de imagen. Por ejemplo, si se captura una imagen de 24 bits con una cámara digital con dimensiones de píxel de 2.048 x 3.072, entonces el tamaño de archivo es igual a (2048 x 3072 x 24) / 8, o 50.331.648 bytes.

#### *Fórmula 2 para el tamaño de archivo*

**Tamaño de archivo = (dimensiones de píxel x profundidad de  
bits) / 8**



## Digitalización

### ❖ Compresión

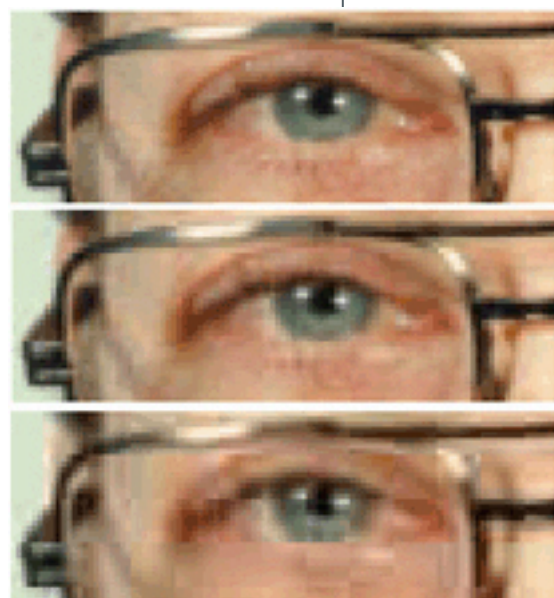
Nombre	CCITT G-4 (ITU-T.6)	JBIG	JPEG	LZW (GIF)	LZW (TIFF, PDF, PS)	ImagePac (PhotoCD)
Estándar/ Patentado	Estándar	Estándar	Estándar	Patentado	Patentado	Patentado
Sin pérdida/ Con pérdida	Sin pérdida	Sin pérdida	Con pérdida	Sin pérdida	Sin pérdida	Con pérdida
Profundidad bits/pixel	1	De 1 a 6	8 ó 24	De 1 a 8	De 1 a 8	24
Resolución múltiple	No	No	No	No	No	Si
Soporte Web	Conexión o aplicación de ayuda	Conexión o aplicación de ayuda	Desde Explorer 2, Navigator 2	Desde Explorer 3, Navigator 2,	Conexión o aplicación de ayuda	Aplicación Java o de ayuda

...y no se olviden de DjVU <http://www.lizardtech.es/demo/>





## Digitalización



100 dpi compresión JPEG baja  
tamaño  
de archivo  
248K

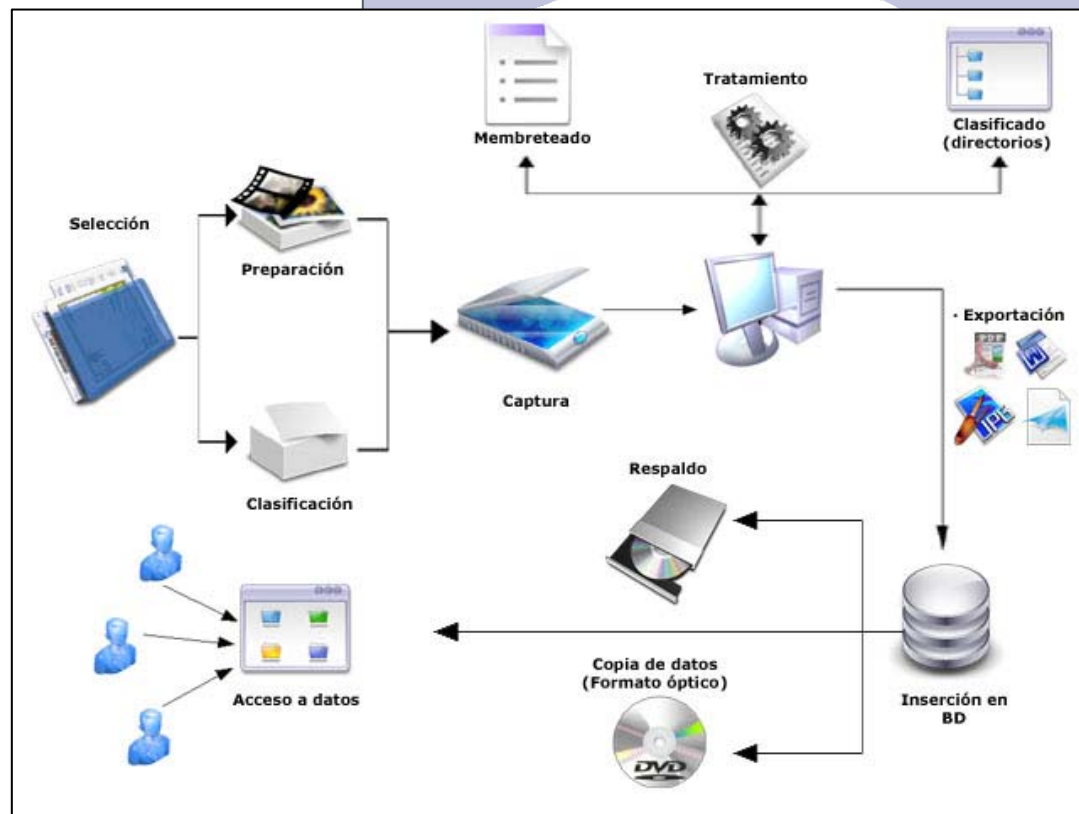
100 dpi compresión JPEG media  
tamaño  
de archivo  
49K

100 dpi compresión JPEG alta  
tamaño  
de archivo  
22K

❖ Resultados según compresión



## Digitalización



- ❖ ¡Comenzar a digitalizar ya! pero...
- ❖ Definir claramente los parámetros, datos y formatos del trabajo
  - ❖ Pruebas preliminares
  - ❖ Formato máster
  - ❖ Formato para difusión
  - ❖ Plazo de realización
  - ❖ Soportes de entrega
  - ❖ Carga de datos
  - ❖ Datos asociados y su carga
- ❖ Ejemplos de pliegos para empresas externas

## Digitalización



- ❖ Buenas prácticas
  - ❖ Emplear formatos habituales (recomendar también para autopublicación)
  - ❖ Cuidado con la presentación
  - ❖ Cuidado con el peso
  - ❖ Sintaxis y estructura de los ficheros
  - ❖ Poner “Copyright” o marcas de agua
  - ❖ Llevar el control de las digitalizaciones mediante una base de datos



## Postproducción

and valid names. Either you already name(s) or the names include characters periods ("."), dashes ("-"), plus signs. the files from an existing release, or name.

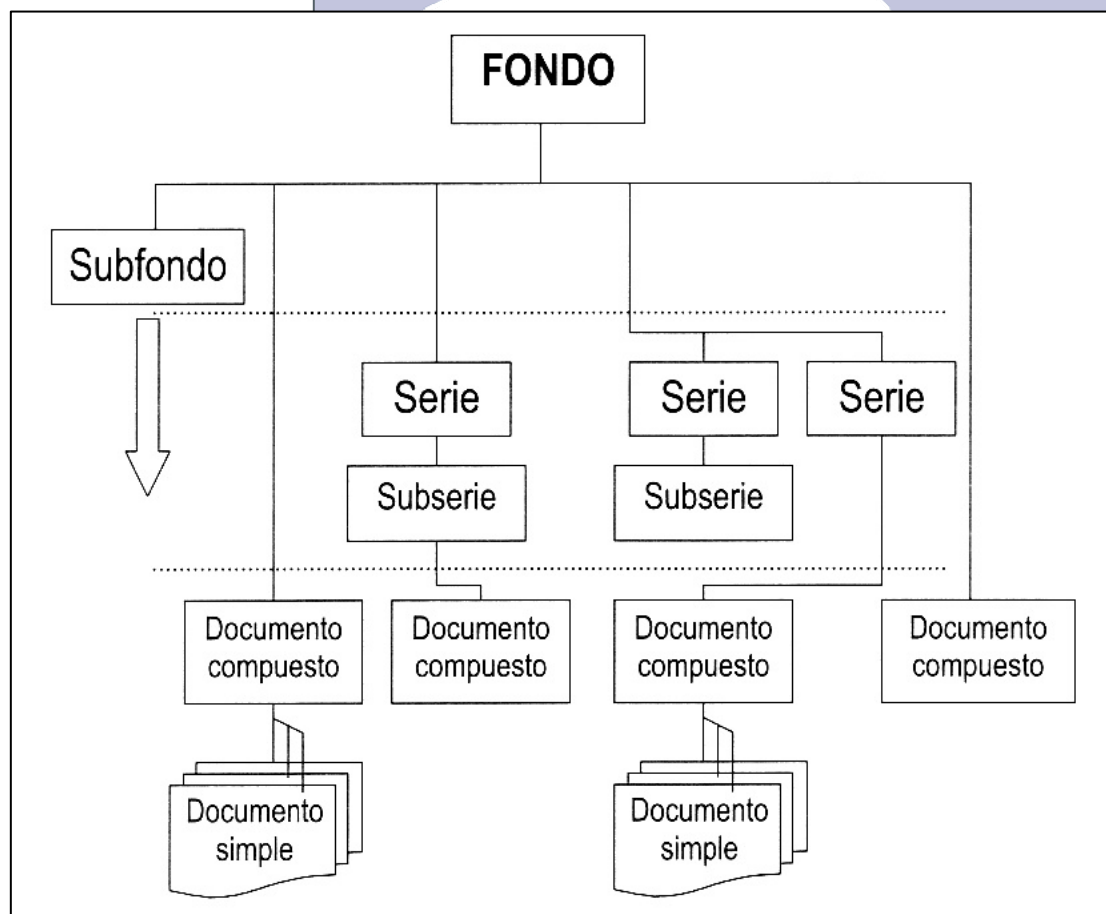
- ☐ TNTConcept-INSTALL\_0.12.6\_en.pdf
- ☐ TNTConcept-INSTALL\_0.12.6\_es.pdf

Add Files and/or Refresh View

- ❖ Aplicación de sintaxis a ficheros máster y secundarios



## Postproducción

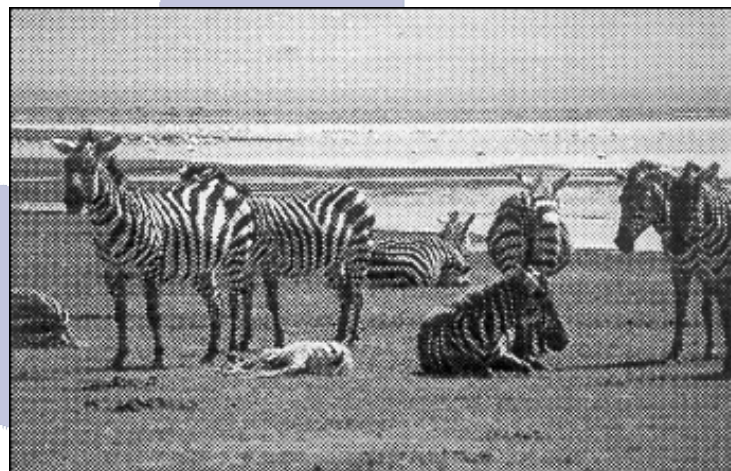


- ❖ Idóneamente la sintaxis y/o la estructura de los ficheros deberá respetar la organización de los originales
- ❖ Cuadros clasificación (ejs.):
  - ❖ AHRSEAPGC  
<http://mdc.ulpgc.es/rseapgc>
  - ❖ AFDJO  
<http://mdc.ulpgc.es/oshanahan>
  - ❖ Voces y ecos  
<http://mdc.ulpgc.es/voces>
  - ❖ ...



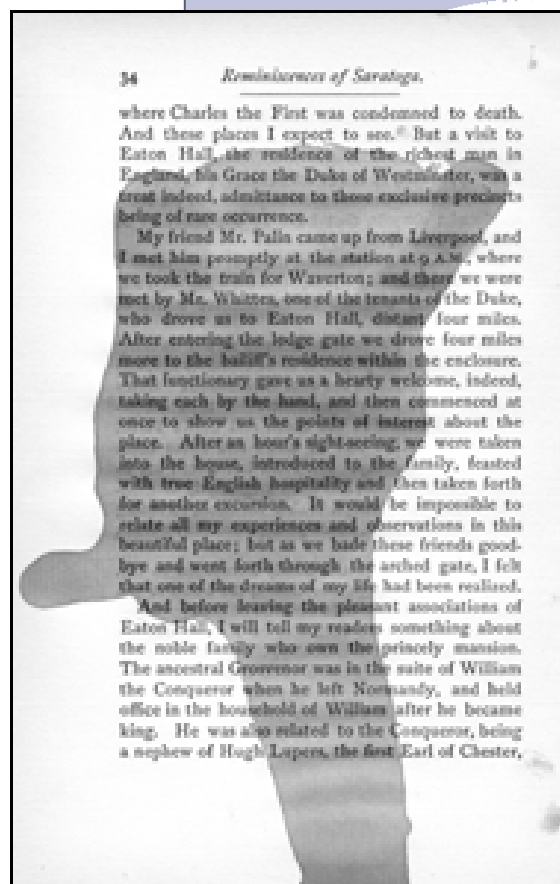
## Postproducción

- ❖ Ajuste del rango dinámico
- ❖ Entramado, efecto “moaré”, etc.



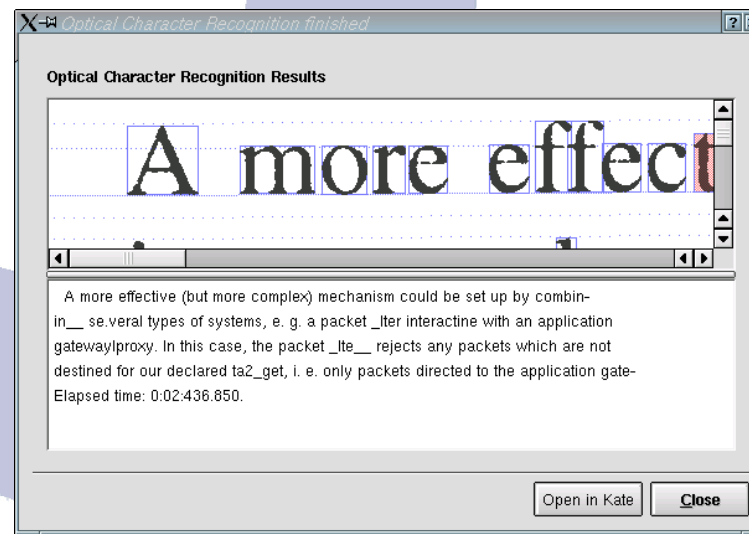
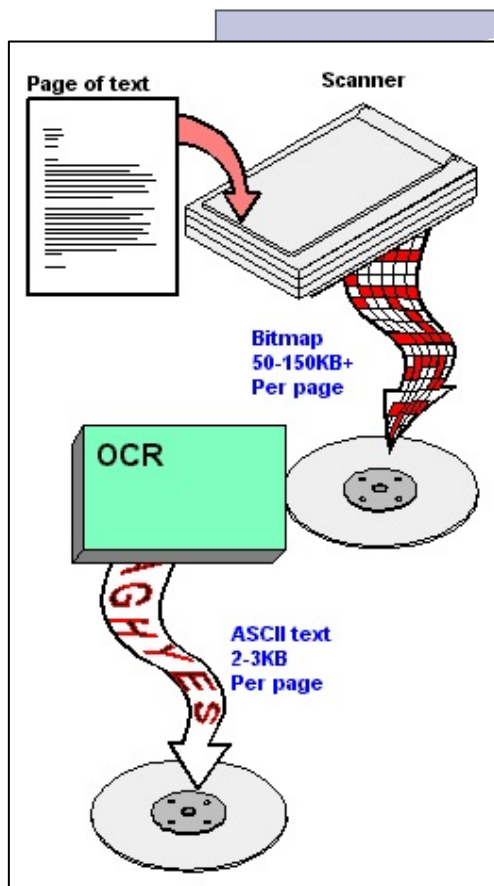
## Postproducción

- ❖ Ajuste correcto del color
- ❖ Trasparencias
- ❖ Manchas

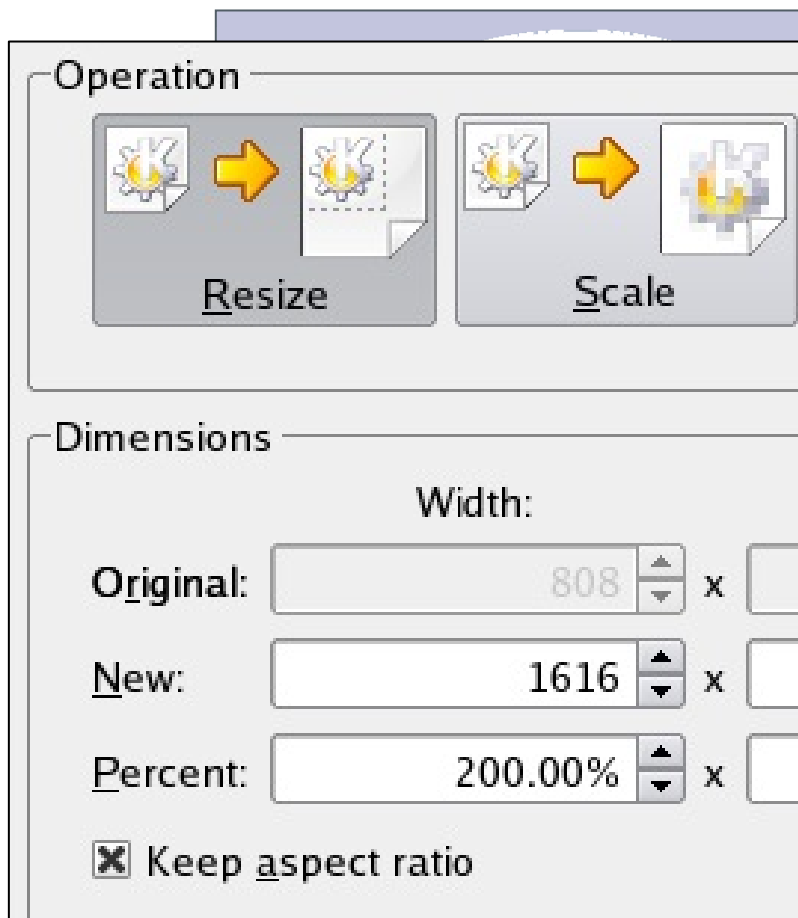


## Postproducción

- ❖ Reconocimiento óptico de caracteres (OCR)



## Postproducción



- ❖ Redimensionamiento
  - ❖ Ficheros máster
  - ❖ 30%, 40%, 50%, etc. del original
  - ❖ Conservando relación de aspecto
  - ❖ Conservando calidad



## Control de calidad



- ❖ Evaluación
  - ❖ Previa
    - ❖ Pruebas y testeo
    - ❖ Envío y recepción de muestras (externalización)
    - ❖ Necesidad de cambiar parámetros o aplicar tratamientos en postproducción
  - ❖ Continua (grandes lotes)
  - ❖ Final
    - ❖ Aleatoria o completa
- ❖ Hardware y software fiable
- ❖ Soportes físicos
  - ❖ Software gratuito de análisis
    - ❖ VSO Inspector
    - ❖ Dvdisaster
    - ❖ CDDoc!







## Metadatos

- ❖ De uso general
  - ❖ Etiquetas META
  - ❖ MARC21XML
  - ❖ Dublin Core (DC)
    - <http://es.wikipedia.org/wiki/DCMI>
    - [http://bvs.sld.cu/revistas/aci/vol14\\_4\\_06/aci09406.htm](http://bvs.sld.cu/revistas/aci/vol14_4_06/aci09406.htm)
    - <http://www.ciw.cl/proyectos/calificadoresDC.html>
    - GENERADOR <http://www.ukoln.ac.uk/metadata/dcdot/>
  - ❖ RDF
- ❖ Específicos
  - ❖ EAD
    - <http://www.hipertext.net/web/pag223.htm>
    - <http://archivo.ayuncordoba.es/EAD/pergaminos/index.html>
  - ❖ METS
    - [http://www.loc.gov/standards/mets/METSOverview\\_spa.html](http://www.loc.gov/standards/mets/METSOverview_spa.html)
  - ❖ METS / ALTO
    - <http://www.ccs-gmbh.com/alto/>
  - ❖ VRA Core
- ❖ Embebidos
  - ❖ PDF
  - ❖ EXIF - IPTC





## Practiquemos con Adobe Acrobat

- ❖ Crear TIFF desde PDF / Crear PDF desde TIFF
- ❖ Crear PDF desde una página en blanco / Desde el escáner
- ❖ Crear carpetas PDF
- ❖ Enlazar parte del texto a una página web
- ❖ Captura de web a PDF
- ❖ Manejar el Organizador
- ❖ Establecer modo de presentación (vista)
- ❖ Poner marcadores
- ❖ Editar metadatos
- ❖ Realizar OCR
- ❖ Enviar por correo electrónico
- ❖ Añadir marca de agua
- ❖ Adjuntar ficheros
- ❖ Insertar firma digital
- ❖ Exportar en diversos formatos
- ❖ Ponerle seguridad





## Preservación digital

*A medida que entramos en la era electrónica de objetos digitales, es importante saber que hay nuevos bárbaros en la frontera y que estamos entrando en una época donde mucho de lo que sabemos hoy, mucho de lo que está codificado y escrito electrónicamente, se perderá para siempre. Estamos, en mi opinión, viviendo en medio de la edad media digital. En consecuencia, igual que los monjes del pasado, le toca a bibliotecarios y archiveros hacerse cargo de la tradición de conservar la historia y el patrimonio documental de nuestros tiempos.*

Terry Kuny  
Developer and administrator of electronic services for IFLA

## Preservación digital







## Preservación digital

- ❖ Grave problemática
  - ❖ Estamos perdiendo nuestra memoria personal (fotos digitales sólo en la cámara, o en un pendrive, o en CD reescribible, o el disco duro...)
  - ❖ Perdemos los testimonios y las habilidades de nuestro entorno y nuestra cultura (patrimonio intangible)
  - ❖ Bibliotecas y archivos ya no son las únicas depositarias de la memoria de la humanidad. Internet es la gran competencia (o la gran ayuda...)





## Preservación digital



- ❖ Grave problemática
  - ❖ Bibliotecarios, archiveros y documentalistas deben asumir un nuevo rol profesional, porque...
  - ❖ ¿Qué hacemos con los documentos que nacen digitales?, ¿cómo controlamos hacernos con ellos?
  - ❖ ¿Qué sucede con lo que los usuarios suben directamente a Internet?
  - ❖ ¿Qué sucede con lo que sólo está en Internet y no en el plano físico?



## Preservación digital



Color photo by Jeff Rotherberg

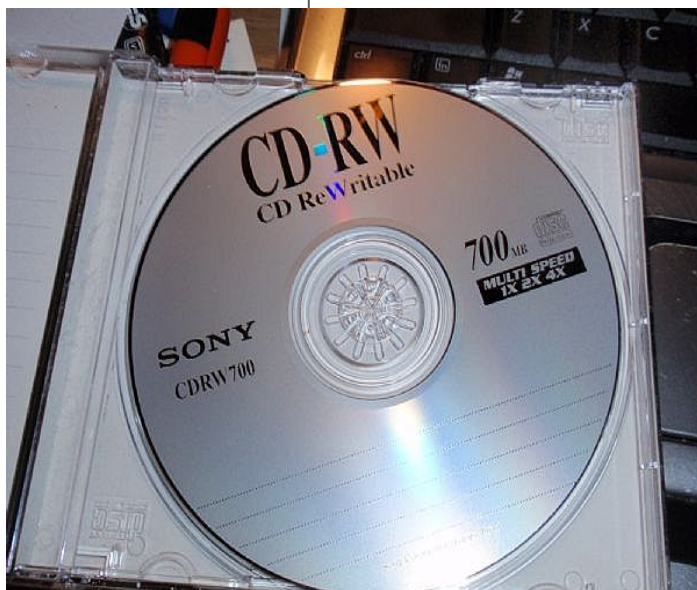
- ❖ Obsolescencia de los originales electrónicos
- ❖ El 93% de la información generada en el último año ha sido de forma electrónica
- ❖ Procedimientos
  - ❖ Utilización de estándares
  - ❖ Migración a nuevos formatos electrónicos
  - ❖ Replicado de soportes físicos
  - ❖ Copias de seguridad
  - ❖ Independencia de los dispositivos
  - ❖ No encriptar
  - ❖ No comprimir
- ❖ Formatos
  - ❖ PDF/A (ISO/CD 19005-1)

<http://en.wikipedia.org/wiki/PDF/A>

<http://www.digitalpreservation.gov/formats/fdd/fdd000125.shtml>

[http://www.gusgsm.com/freguntas\\_respuestas\\_frecuentes\\_faq\\_estandar\\_pdf](http://www.gusgsm.com/freguntas_respuestas_frecuentes_faq_estandar_pdf)

## Preservación digital



- ❖ Uno de los sistemas ideales
  - ❖ Discos ópticos (sólo en primera instancia, buenas marcas, duplicados, separados, bien almacenados (magnetismo, fuego, humedad...))
  - ❖ Discos duros magnéticos (buenas marcas, duplicados, separados, bien almacenados (magnetismo, fuego, humedad, temperatura...))
- ❖ Servidores
  - ❖ RAID 5
  - ❖ Copia en backup
- ❖ Servidores virtuales y almacenamiento masivo en línea SAN/NAS balanceado
- ❖ Doble centro de cálculo (“alta disponibilidad”)



## Preservación digital



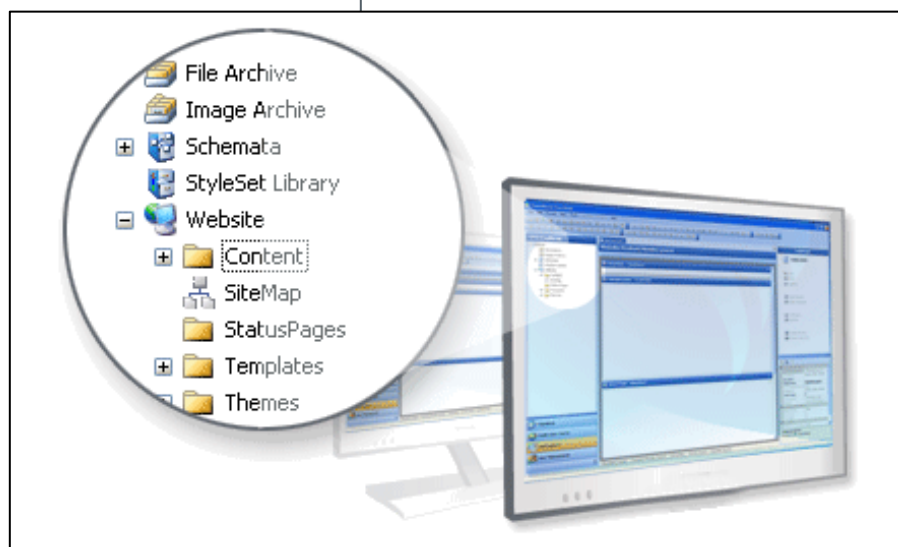
- ❖ Metadatos y protocolos
  - ❖ Información sobre el contenido
  - ❖ PREMIS  
<http://en.wikipedia.org/wiki/PREMIS>
  - ❖ PURL y URI  
<http://purl.org>
- ❖ Programa *Memoria del Mundo* (UNESCO)  
[http://es.wikipedia.org/wiki/Programa\\_Memoria\\_del\\_Mundo](http://es.wikipedia.org/wiki/Programa_Memoria_del_Mundo)



## Difusión: accesibilidad y visibilidad

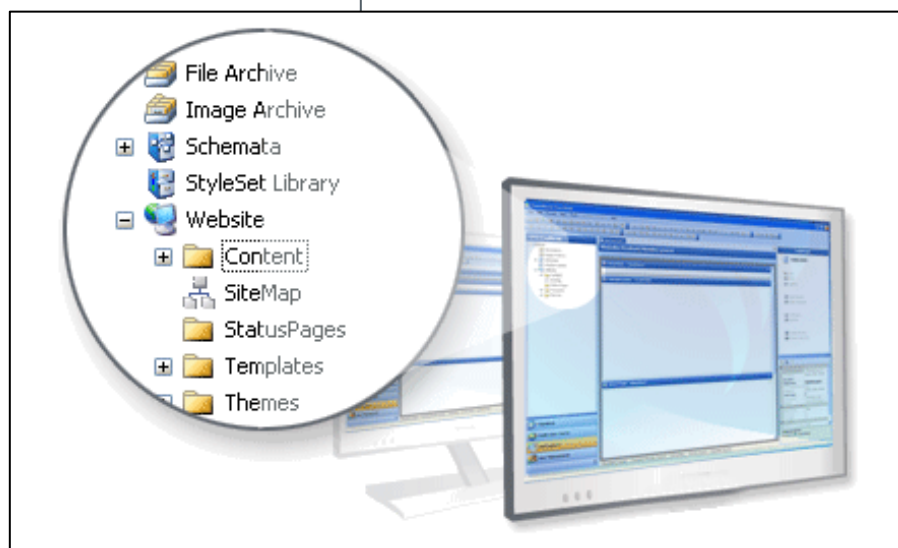
### ❖ Software

- ❖ CDS Invenio <http://cdsware.cern.ch/invenio/index.html>  
Ejemplos: <http://ddd.uab.es>
- ❖ CONTENTdm <http://www.contentdm.com>  
Ejemplos: <http://mdc.ulpgc.es> <http://repositorio.ulpgc.es>
- ❖ Digitool <http://www.greendata.es/index.php?option=content&task=view&id=35>  
Ejemplos: <http://bibotecadigitalhispanica.bne.es/R/>
- ❖ DSpace <http://www.dspace.org>  
Ejemplos: <http://digital.csic.es>
- ❖ E-Prints <http://www.eprints.org/software/>  
Ejemplos: <http://eprints.ucm.es>
- ❖ Greenstone <http://www.greenstone.org>  
Ejemplos: <http://ibdigital.uib.es/gsd/cgi-bin/library>
- ❖ ... o desarrollos propios.  
Ejemplos: Memoria digital de Canarias (mdC) 2da. ed.  
<http://alojamiento.ulpgc.es/mdc>





## Difusión: accesibilidad y visibilidad



### ❖ Lo que utilizamos en la BULPGC:

#### ❖ Software comercial (CONTENTdm)

<http://www.contentdm.com>

Visible en: <http://mdc.ulpgc.es>

#### ❖ Software comercial... "sin coste" (PANDORA)

Visible en: <http://jable.ulpgc.es>

#### ❖ Encargo de desarrollos personalizados (BUStreaming)

Visible en: <http://bustreaming.ulpgc.es>

#### ❖ Software libre (Dspace)

<http://www.dspace.org>

Visible en: <http://acceda.ulpgc.es>

## Difusión: accesibilidad y visibilidad



- ❖ Accesibilidad
  - ❖ WAI (Test T.A.W.)
  - ❖ US Section 508
  - ❖ Adaptación a dispositivos móviles
- ❖ Visibilidad
  - ❖ Identificación (HTML "Title")
  - ❖ Etiquetado (META) + DCMI
  - ❖ Posicionamiento en buscadores
  - ❖ Protocolos como OAI-PMH



## Difusión: usabilidad

- ❖ Usabilidad
  - ❖ Presentación e información general
  - ❖ Propósitos y cooperación
  - ❖ Ayuda con instrucciones técnicas
  - ❖ Ofrecer la descarga de las herramientas necesarias para la consulta (plugins, etc.)
  - ❖ Navegabilidad
  - ❖ Datos de contacto



## Difusión: marketing



- ❖ Diseño atractivo para nuestras digitalizaciones
- ❖ Impresión de folletos, trípticos, cartelería, etc. que también se ponen en digital
- ❖ Mailing (electrónico –personal y a listas de distribución- y postal)
- ❖ Presentaciones públicas
- ❖ Presentaciones en congresos del sector
- ❖ Generar noticias para la prensa
- ❖ Coincidir con efemérides
- ❖ Entrevistas en radio y televisión
- ❖ Redacción de artículos para publicaciones especializadas y también generales
- ❖ Llamada pública a la cooperación (y agradecimiento posterior)





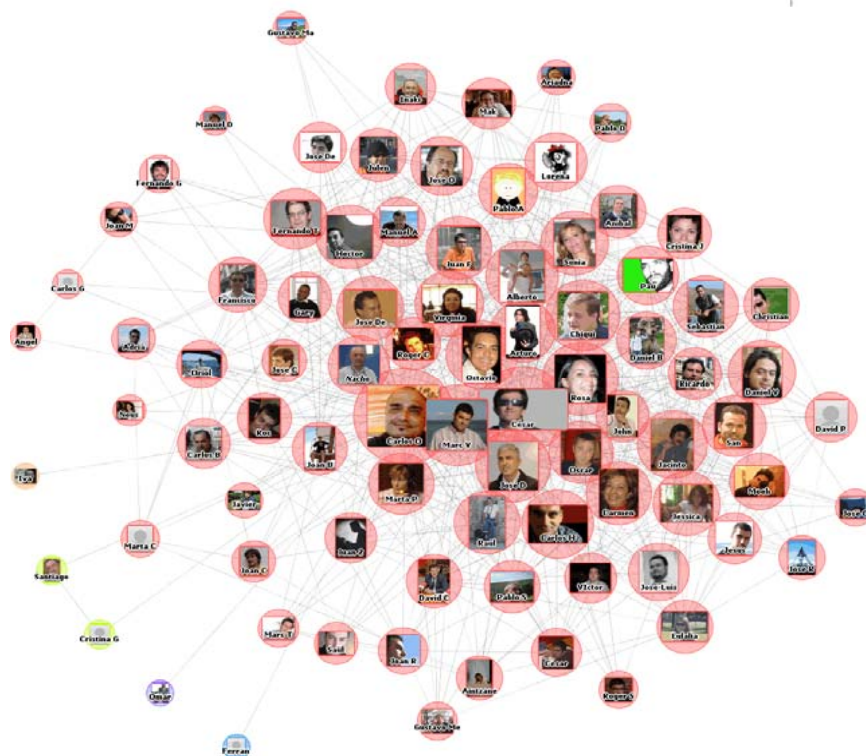
## Dónde deben “vernós” (sólo si somos *Open Access*)

- ❖ [ROAR](#)  
(Registry of Open Access Repositories)
- ❖ [OpenDOAR](#)  
(Directory of Open Access Repositories)
- ❖ [Oaister](#)
- ❖ [WorldCat](#) (OCLC)
- ❖ [Recolecta](#)
- ❖ [Hispana](#)
- ❖ [Busca repositorios](#) (CSIC)
- ❖ [TDR / TDX](#)
- ❖ .....





## Usuarios



*Internet es la mayor biblioteca que existe. Google, su mayor bibliotecario, el que encuentra todas las páginas y todas las referencias; pero Google pertenece a un Internet antiguo; el de ahora es más móvil, más social, donde la gente no sólo lee, sino que habla, escribe, fotografía o filma. Es otra forma de usar Internet. Internet ha pasado de ser la gran biblioteca a ser la gran conversación.*

David Sifry, fundador de [Technorati](#), en una entrevista en *El País*.

## Usuarios



La **Web 2.0** son aplicaciones y servicios de Internet que utilizan la inteligencia colectiva para proporcionar recursos interactivos en red dando al usuario el control.

Los **Usuarios 2.0** son nativos digitales. Nacieron en un mundo con ordenadores e interactúan cotidianamente con nuevos medios. Los móviles, videojuegos, email, chat e Internet son parte integral de sus vidas

La **Biblioteca 2.0** es convertir nuestro espacio físico o digital en un espacio interactivo, retroalimentado, capaz de no sólo de atender las necesidades de nuestros usuarios sino de hacer que se sientan parte de nuestra propia Biblioteca.



## Usuarios



### ❖ Herramientas y servicios

- ❖ Novedades
- ❖ Alertas
- ❖ Comentarios
- ❖ Evaluación de contenidos
- ❖ Formatos para el mismo contenido
- ❖ Blogs
- ❖ RSS
- ❖ Redes sociales
- ❖ Responder siempre



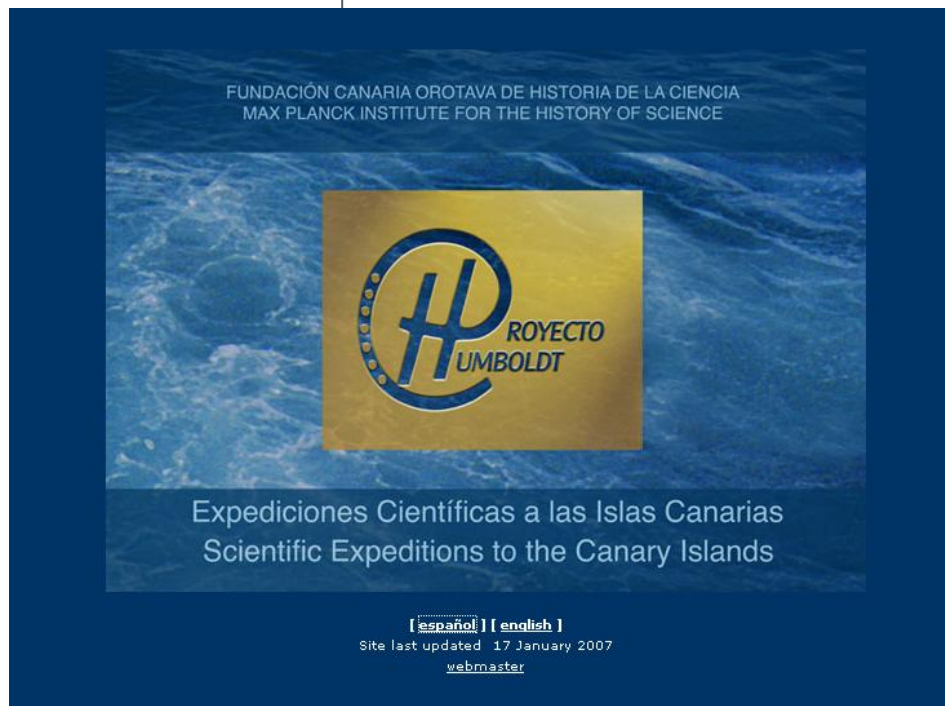
## Portales canarios que hay que conocer



## Memoria digital de Lanzarote



## Portales canarios que hay que conocer



## Proyecto Humboldt Expediciones científicas a las Islas Canarias

<http://humboldt.mpiwg-berlin.mpg.de>



## Portales canarios que hay que conocer

### Memoria digital de Canarias (mdC)

- ❖ Abierta a la cooperación puntual o general para la digitalización de fondos de entidades y particulares
- ❖ Cumplimiento de la normativa y protocolos de repositorios abiertos
- ❖ Ramificada en subportales con entidad propia de todas las temáticas, épocas y soportes



## Portales canarios que hay que conocer

### MDC. Renovarse o morir...

#### ❖ Antes

- ❖ Descripción ad-hoc
- ❖ DJVU
- ❖ Realmedia
- ❖ PDF
- ❖ No OAI

#### ❖ Ahora

- ❖ Descripción Dublin Core
- ❖ JPEG
- ❖ Flash
- ❖ PDF
- ❖ OAI-PMH



## Portales canarios que hay que conocer

Los beneficios de la cooperación...



*InfoNorteDigital.com*  
Información del Norte de Gran Canaria



Reserva de la Biosfera



**Casa de Colón**



## Portales canarios que hay que conocer

Memoria digital de Canarias (mdC) se amplía para intentar  
 dar cabida y difundir todo el patrimonio documental canario



## Portales canarios que hay que conocer



### JABLE Archivo de series digitalizadas

- ❖ Digitalizando prensa canaria corriente desde 1996
- ❖ También prensa histórica
- ❖ Incluye boletines, revistas, etc.
- ❖ Previsto 221 cabeceras
- ❖ También pub. peninsulares y coloniales africanas
- ❖ 6 millones páginas y sumando
- ❖ Buscador múltiple: textual, cronológico y geográfico

❖ Más info en <http://blip.tv/file/1408241/>



## Portales canarios que hay que conocer

### ACCEDA

#### Documentación científica de la ULPGC en abierto

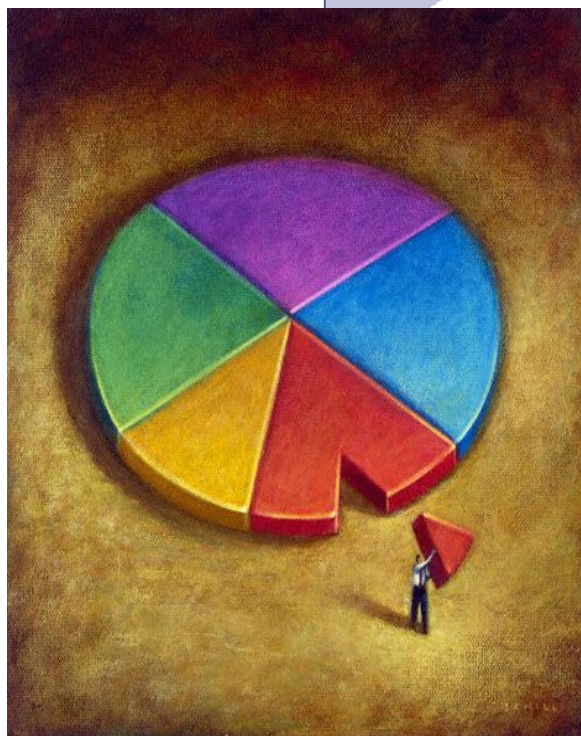


- ❖ Repositorio institucional que recogerá toda la producción académica, investigadora y administrativa de la ULPGC
- ❖ Los usuarios se autopublican directamente
- ❖ Licencias abiertas *Creative Commons*
- ❖ Tesis, tesinas, proyectos fin de carrera, memorias de máster, informes, artículos, comunicaciones, ponencias, jornadas, actas, objetos de aprendizaje, etc.
- ❖ Autoridades recomendaciones FECYT
- ❖ Indización nomenclatura UNESCO con búsqueda ontológica
- ❖ DCMI cualificado + OAI-PMH
- ❖ Sistema de preservación digital incorporado
- ❖ RefWorks + SFX + Recomendar ítems + Blog + Web 2.0

## La Biblioteca Universitaria de la ULPGC. Lo que tenemos... y lo que nos queda por hacer



- ❖ mdC
- ❖ Jable
- ❖ Acceda
- ❖ BUStreaming
- ❖ Urchin
- .....
- ❖ Congreso y consorcio regional
- ❖ Harvester OAI
- ❖ Consigna
- ❖ Mirroring
- ❖ Generar nuevos contenidos propios



## BULPGC. Estadísticas de digitalización

	Memoria digital de Canarias (mdC)	Prensa canaria / Jable	Repositorio académico / Acceda
2004	2.673.809 26.855	-----	4.544.269 73.131
2005	5.195.584 101.231	-----	6.795.886 464.663
2006	-----		
2007	26.576.253 1.744.179		
2008	26.556.153 709.803	2.831.265 17.978	43.360 2.990

Cifras en negro: número de consultas

Cifras en azul: número de sesiones (visitantes)



## Portales nacionales e internacionales que hay que conocer



- ❖ [Biblioteca Virtual Miguel de Cervantes](#)  
(Univ. Alicante+Banco Santander)
- ❖ [Biblioteca Digital Hispánica](#) (BNE)
- ❖ [Biblioteca virtual del Patrimonio Bibliográfico](#)  
(Ministerio de Cultura)
- ❖ [Biblioteca virtual de Prensa Histórica](#) (MinCul.)
- ❖ [Hispana](#) (Ministerio de Cultura)
- ❖ [Pares. Portal de archivos españoles](#) (MinCul.)
- ❖ [Recolecta](#)
- ❖ [Busca repositorios](#) (CSIC)
- ❖ [TDR / TDX](#) (CBUC)
- ❖ .....
- ❖ [Google Libros](#)
- ❖ [Google Académico](#)
- ❖ [Europeana](#) (UE)
- ❖ [DRIVER](#) (UE)







## Recursos y documentación

- ❖ Guía para la digitalización de documentos impresos (Universidad Nacional de la Plata)  
[http://163.10.30.7:9673/memoria/guia\\_digitalizacion\\_fahce.pdf/vi ew](http://163.10.30.7:9673/memoria/guia_digitalizacion_fahce.pdf/vi ew)
- ❖ Tutorial de digitalización de la Universidad de Cornell (USA)  
<http://www.library.cornell.edu/preservation/tutorial-spanish/contents.html>
- ❖ Tutorial La digitalización y las bibliotecas digitales (FAO & UNESCO iMARK)  
[http://www.imarkgroup.org/moduledescrE\\_es.asp](http://www.imarkgroup.org/moduledescrE_es.asp)





## Recursos y documentación



- ❖ ALONSO FERNÁNDEZ, Juan. Digitalización, catalogación y recuperación de información en los archivos fotográficos: un estado de la cuestión. Barcelona: UAM, Facultad de Ciencias de la Comunicación, 2007. En línea [pulsando aquí](#).
- ❖ ARQUERO AVILÉS, Rosario y GARCÍA-OCHOA ROLDÁN, María Luisa. *La hemeroteca de prensa*. Gijón: Trea, 2005.
- ❖ BORRAS, Antoni y otros. *Programari de codi lliure per gestionar dipòsits digitals: el procés de tria dut a terme al CBUC*. En: BiD. Textos universitaris de biblioteconomia i documentació. 2006, n. 16. En línea [pulsando aquí](#).
- ❖ CAMPUZANO JUAN, Francisco y GIMÉNEZ CHORNET, Vicent. Recomendaciones técnicas para la captura y digitalización de fondos de archivos. Valencia: Arxiu del Regne de València, 2005. En línea [pulsando aquí](#).
- ❖ *Directrices para la preservación del patrimonio digital*. Canberra: UNESCO, 2003. En línea [pulsando aquí](#).
- ❖ Directrices para proyectos de digitalización de colecciones y fondos de dominio público, en particular para aquellos custodiados en bibliotecas y archivos. Traducción española. Madrid: Ministerio de Cultura, 2005. En línea [pulsando aquí](#).
- ❖ *Directrius per a projectes de digitalització de col·leccions i fons de domini públic, en especial els de biblioteques i arxius*. Barcelona: Col·legi de Bibliotecaris-Documentalistes de Catalunya, 2006. En línea [pulsando aquí](#).



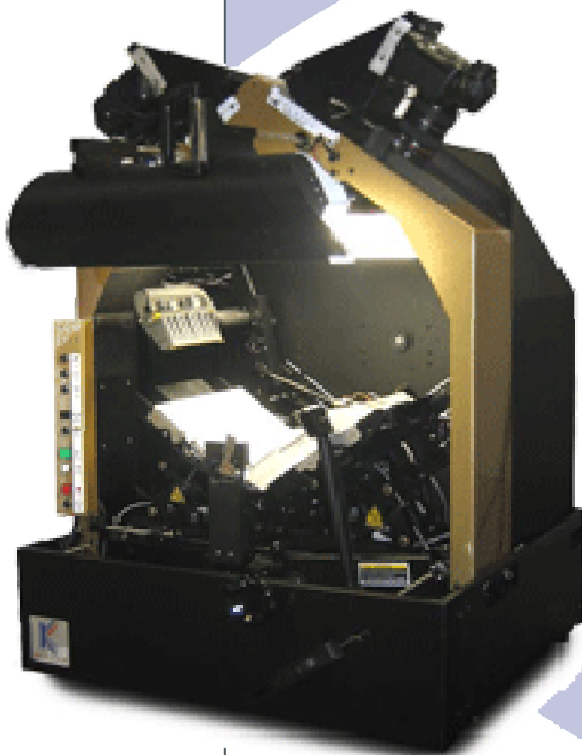
## Recursos y documentación



- ❖ HERNÁNDEZ PÉREZ, Tony y otros. *Open Access: el papel de las bibliotecas en los repositorios institucionales de acceso abierto*. En Anales de documentación. 2007, n. 10. En línea [pulsando aquí](#).
- ❖ KEEFER, Alice y GALLART, Núria. *La preservación de recursos digitales: el reto para las bibliotecas del siglo XXI*. Barcelona: UOC, 2007. En línea [pulsando aquí](#).
- ❖ ORDEN CUL/1014/2007, de 30 de marzo, por la que se constituye la Comisión Española sobre la digitalización y la accesibilidad en línea del material cultural y la conservación digital.
- ❖ ORDEN JUS/1468/2007, de 17 de mayo, sobre impulso a la informatización de los registros civiles y la digitalización de sus archivos.
- ❖ RESOLUCIÓN de 24 de octubre de 2007, de la Agencia Estatal de Administración Tributaria, sobre procedimiento para la homologación de software de digitalización contemplado en la Orden EHA/962/2007, de 10 de abril de 2007.
- ❖ SALVADOR BENÍTEZ, Antonia y RUIZ RODRÍGUEZ, Antonio A. *Archivos fotográficos: pautas para su integración en el entorno digital*. Granada: Universidad, 2006.
- ❖ SCHREIBMAN, Susan (ed.) y otros. *Best practices guidelines for digital collections*. University of Maryland Libraries, 2007. En línea [pulsando aquí](#).
- ❖ VIVES, Josep (coord.). *Digitalización del patrimonio: archivos, bibliotecas y museos en la red*. Barcelona: UOC, 2009. En línea [pulsando aquí](#).



## Recursos y documentación



- ❖ Vídeos de escáners de alta producción y más...
  - ❖ Canon <http://www.youtube.com/watch?v=4FtHFXfohoY>
  - ❖ SMA <http://www.youtube.com/watch?v=1-OfcoqL7Qs>
  - ❖ Zeutschel <http://www.youtube.com/watch?v=A9vOnU0mdbw>
  - ❖ BookDrive DIY <http://www.youtube.com/watch?v=BfVvfRARuts&feature=related>
  - ❖ Atiz <http://www.youtube.com/watch?v=XheLKg3FVVQ&feature=related>
  - ❖ ScanRobot <http://www.youtube.com/watch?v=y16rNqnxj0U&feature=related>
  - ❖ Kirtas <http://www.youtube.com/watch?v=nVIR8HHUslg&feature=related>







**Y en  
cualquier  
caso buen  
humor...**

**Gracias por asistir  
y si necesitan  
algo más:**

vmacias@pas.ulpgc.es  
Tlf. 928458680



<http://biblioteca.ulpgc.es>

email [sdd@ulpgc.es](mailto:sdd@ulpgc.es)  
tlf 928457249